DOCUMENT RESUME

ED 444 198 CS 510 376

AUTHOR Madsen, Thomas O.
TITLE Sproglyde: fysiske kontinua og mentale kategorier (Language Sounds: Physical Continua and Mental Categories). Odense Working Papers in Language and Communication, No. 20.
INSTITUTION Odense Univ. (Denmark). Inst. of Language and Communication.
PUB DATE 2000-06-00
NOTE 181p.
AVAILABLE FROM Institute of Language and Communication, University of Southern Denmark, Main Campus: Odense University, Campusvej 55, DK-5230 Odense M., Denmark. Email: lfn@language.sdu.dk.
PUB TYPE Reports - Research (143)
LANGUAGE Danish
EDRS PRICE MF01/PC08 Plus Postage.
DESCRIPTORS *Auditory Perception; Brain Hemisphere Functions; Communication Research; Higher Education; *Language Processing; Language Research; Psycholinguistics
IDENTIFIERS *Categorical Perception; *Dichotic Listening

ABSTRACT

This study presents an empirical investigation of basic processes in the perception of speech sounds. The experimental methods applied have their roots in two different psycholinguistic research paradigms; i.e., "categorical perception" and "dichotic listening." In the categorical perception paradigm, listeners' categorization of speech sounds is compared with their ability to discriminate speech-like stimuli forming a physical continuum (such as the range between a T-sound and a D-sound). In the dichotic listening paradigm, differences in listeners' identification of speech stimuli presented to either left or right ear are interpreted as the result of hemispheric lateralization. Extensive results include the following: (1) the acoustic consonant-continuum revealed a high degree of categorical effect; (2) when the same acoustic continuum was presented in the dichotic listening experiment, subjects showed a markedly better ability to identify stimuli presented to the right ear than stimuli presented to the left ear; (3) on the other hand, the acoustic vowel-continuum revealed no categorical effect at all; and (4) similarly, no general dichotic effect was found when the same vowel-continuum was presented in the dichotic listening experiment. Findings suggest that there exists a clear connection between the categorical and the dichotic effect. (Contains 211 references and 51 figures of data. An appendix contains acoustic analyses of synthetic vowel-continua.) (SR)

Reproductions supplied by EDRS are the best that can be made
from the original document.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

T. Madsen

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

UNIVERSITY OF SOUTHERN DENMARK

ED 444 198

Sproglyde: fysiske kontinua og mentale kategorier

Thomas O. Madsen

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☐ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

CS 510 376

No. 20
June
2000

ODENSE WORKING PAPERS IN LANGUAGE AND COMMUNICATION

BEST COPY AVAILABLE

Thomas O. Madsen

Sproglyde: fysiske kontinua og mentale kategorier

Odense Working Papers in Language and Communication
No 20, June 2000

General editor: Steffen Nordahl Lund (snl@language.sdu.dk)

Institute of Language and Communication
University of Southern Denmark,
Main Campus: Odense University
Campusvej 55
DK-5230 Odense M.
(lfn@language.sdu.dk)

Printed in Denmark by Odense Universitets Trykkeri

Indholdsfortegnelse

Forkortelsesliste		ii
Forord		iii
Resumé		v
1.1	Tegnsystem, kode og medium	1
1.2	Akustisk segmentering og invariante træk	6
1.3	Teorier om sprogperception	20
2.1	Introduktion til kategorisk perception	31
2.2	Forsøg med kategorisk perception	48
2.3	Resultater fra forsøg med kategorisk perception: et S-T-D-kontinuum	76
2.4	Resultater fra forsøg med kategorisk perception: et I-E-Æ-A-kontinuum	91
3.1	Introduktion til dikotisk lytning	109
3.2	Forsøg med dikotisk lytning	117
3.3	Resultater fra forsøg med dikotisk lytning: et S-T-D-kontinuum	128
3.4	Resultater fra forsøg med dikotisk lytning: et I-E-Æ-A-kontinuum	136
3.5	Resultater fra forsøg med dikotisk lytning: [la], [ma], [na], [fa] og [va]	146
4	Konklusion	150
Litteraturliste		155
Bilag A		163

Forkortelsesliste

CV-stavelser	konsonant-vokal-stavelser
DL	dikotisk lytning
EEG	elektroencefalogram
F0	grundtone
F1	formant 1
F2	formant 2
F3	formant 3
FFT	Fast Fourier Transform
fMRI	functional magnetic resonance imaging
fp	forsøgsperson(er)
HAS	High-Amplitude Sucking
HR	Heart Rate
Hz (eller cps)	antal svingninger per sekund
HØ	højre øre
KP	kategorisk perception
ms	milisekund(er)
PET	positron emission tomography
SLI	specific language impairment
VOT	voice onset time
VØ	venstre øre

Forord

Nærværende *Odense Working Papers in Language and Communication* er i store træk identisk med specialeafhandlingen *Kontinuitet og diskontinuitet. En eksperimentalfonetisk undersøgelse af lateraliserede perceptionsprocesser*, som jeg indleverede ved SDU-Odense Universitet i efteråret 1998.

Rækken af forsøg, der er beskrevet i *Sproglyde: fysiske kontinua og mentale kategorier*, blev alle udført i løbet af forårs- og sommermånederne 1998. Jeg er taknemmelig over for de mange personer, der – mere eller mindre frivilligt – har deltaget i forsøgene, samt IT-funktionen, Humaniora, der stillede teknisk udstyr og ekspertise til rådighed.

Kirsten Gregersen – fra daværende Center for Anvendt Lingvistik og Fremmedsprogspædagogik – var min vejleder i forbindelse med speciale-afhandlingen, og jeg skylder hende tak for stort engagement, utallige lærerige kommentarer og nyttige råd.

I forbindelse med udarbejdelsen af *Sproglyde: fysiske kontinua og mentale kategorier* er jeg professor Hans Basbøll meget taknemmelig for mange stilistiske og indholdsmæssige kommentarer, der har forbedret manusskriptet betydeligt.

Endelig skal der lyde en tak til Institut for Sprog og Kommunikation for at give mig mulighed for at få *Sproglyde: fysiske kontinua og mentale kategorier* udgivet.

Resumé

Formålet med nærværende studium er at foretage en empirisk funderet undersøgelse af grundlæggende processer i perceptionen af sproglyde. De anvendte forsøgsmetoder tager udgangspunkt i de eksperimentalfonetiske/psykolingvistiske paradigmer 'kategorisk perception' og 'dikotisk lytning'.

Et typisk forsøg med kategorisk perception består af to dele: en identifikationsøvelse og en diskriminationsøvelse. I identifikationsøvelsen skal forsøgspersonerne identificere de præsenterede stimuli, hvorimod de i diskriminationsøvelsen skal forsøge at skelne mellem stimuli, der præsenteres i par. De præsenterede stimuli udgør almindeligvis et akustisk kontinuum med varierende VOT (voice onset time) eller formantfrekvenser. Hensigten med forsøget er at afsløre forsøgspersonernes evne til at skelne mellem akustisk lige forskellige stimuli, som de har tilskrevet hhv. samme og forskellige sproglige kategorier. Hvis de grundlæggende processer i perceptionen af sproglyde var baseret på akustisk information, kunne man forvente, at forsøgspersonerne skelnede lige godt mellem stimuli, de havde tilskrevet hhv. samme og forskellige kategorier. Imidlertid viser forsøg med præsentation af klusiler, at det er markant sværere at skelne mellem stimuli, der kategoriseres ens, end stimuli, der kategoriseres forskelligt. Dette fænomen betegnes almindeligvis 'den kategoriske effekt' og fortolkes som resultat af fonetisk perception baseret på diskrete artikulatoriske kommandoer. Det er dog ikke alle sproglyde, der fremkalder en (lige stor) kategorisk effekt. Gennem de sidste 45 år har en lang række forsøg med kategorisk perception vist, at lukkelyde fremkalder den kraftigste kategoriske effekt, nasaler, lateraler og frikativer en mindre kategorisk effekt, og at vokaler perciperes kontinuerligt.

I et andet forskningsparadigme, dikotisk lytning, har man ligeledes kunnet konstatere en markant forskel mellem perception af klusiler, perception af lateraler og frikativer, og perception af vokaler. I forsøg med dikotisk lytning, dvs. dobbeltsidig hørepåvirkning, præsenteres forsøgspersonerne simultant for to stimuli – én i hvert øre – hvorefter de skal identificere én af de pågældende lyde. Umiddelbart skulle man forvente, at forsøgspersonerne identificerede fokuslydene lige ofte, uanset om de blev præsenteret for højre eller venstre øre. Imidlertid viser de fleste forsøg med dikotisk præsentation af klusiler en præference for højre øre, dvs. en signifikant bedre identifikation af stimuli præsenteret for højre øre end af stimuli præsenteret for venstre øre. Dette fænomen betegnes 'den dikotiske effekt' og fortolkes almindeligvis som udtryk for hemisfæreasymmetri, nemlig venstre hjernehalvdels dominans for sprog. Heller ikke i forsøg med dikotisk lytning fremkalder alle sproglyde en lige stor effekt. Resultaterne af de fleste forsøg er, at klusiler fremkalder den største dikotiske effekt,

lateraler og frikativer en mindre dikotisk effekt, og at vokaler ikke fremkalder nogen dikotisk effekt.

I begge eksperimentalfonetiske/psykolingvistiske paradigmer arbejdes der ud fra en hypotese om, at perceptionen af sprog involverer et særligt sprogmodul, en såkaldt sprogprocessor, der varetager analysen af artikulatorisk komplekse sproglyde. I den ene forskningstradition underbygges hypotesen af forsøg, der viser, at klusiler perciperes som diskrete størrelser og vokaler mere kontinuerligt. I den anden forskningstradition anses hypotesen for at være den eneste plausible forklaring på, at forsøg med dikotisk lytning afslører en hemisfæreasymmetri ved præsentation af klusiler, men ikke ved vokallyde. I begge forskningsparadigmer har man dog hidtil kun gjort sig teoretiske overvejelser om en eventuel sammenhæng mellem den kategoriske og den dikotiske effekt.

For at fastslå om der eksisterer en sammenhæng mellem den kategoriske og den dikotiske effekt, udarbejdede jeg et VOT-kontinuum og et formantkontinuum, der skulle udgøre hhv. konsonant- og vokalstimuli i to forskellige forsøg, et med kategorisk perception og et andet med dikotisk lytning. Til forsøgene udviklede jeg et computerprogram, DikoTest 3.0, specielt konstrueret til kontrolleret præsentation af stimuli og registrering af forsøgspersonernes respons. Gruppen af forsøgspersoner, der blev anvendt i begge forsøg, bestod af 57 personer fordelt på 41 danskere (28 kvinder og 13 mænd) og 16 finner (14 kvinder og 2 mænd) i alderen 12-61 år.

I forsøget med kategorisk perception afslørede det akustiske konsonantkontinuum en høj grad af kategorisk effekt: I identifikationsforsøget inddelte forsøgspersonerne det akustiske kontinuum i skarpt afgrænsede kategorier; i diskriminationsforsøget havde forsøgspersonerne meget svært ved at skelne mellem stimuli, som de selv havde tilskrevet samme kategori, hvorimod de meget lettere kunne skelne mellem stimuli fra forskellige kategorier. Da samme akustiske kontinuum blev præsenteret i forsøget med dikotisk lytning, viste forsøgspersonerne en markant bedre evne til at identificere stimuli præsenteret for højre øre end stimuli præsenteret for venstre øre. Omvendt fremkaldte det akustiske vokalkontinuum slet ingen kategorisk effekt: Forsøgspersonerne var i stand til relativt konsekvent at tilskrive stimuli de samme kategorier, og i diskriminationsøvelsen gjorde det ingen forskel for forsøgspersonerne, om de skulle skelne mellem stimuli fra samme eller forskellige kategorier
– udfaldet var lige ringe. Ligeledes fandt jeg heller ikke nogen generel dikotisk effekt, da jeg præsenterede det samme vokalkontinuum i forsøget med dikotisk lytning: forsøgspersonerne var generelt ikke bedre til at identificere stimuli præsenteret for højre øre end stimuli præsenteret for venstre øre.

Da flere forsøg har vist en tendens til en mindre markant kategorisk og dikotisk effekt ved præsentation af bl.a. lateraler og frikativer, valgte jeg endvidere at inddrage CV-stavelserne [la], [ma], [na], [fa] og [va]. Disse stimuli blev præsenteret i forsøget med dikotisk lytning og afslørede – som forventet – en dikotisk effekt, dog mindre end den jeg fandt ved præsentation af det akustiske VOT-kontinuum. På baggrund af resultaterne fra omtalte forsøg kan jeg konkludere, at der eksisterer en klar sammenhæng mellem den kategoriske og den dikotiske effekt.

Hensigten med at inddrage en gruppe finske forsøgspersoner i undersøgelsen var at finde ud af, om forsøgspersoner, der taler forskellige modersmål, kategoriserer de samme akustiske kontinua ens eller forskelligt. Hvis den kategoriske effekt skyldes sproglydenes akustiske struktur, skulle man forvente, at den determinerer de fonetiske kategorier – dvs. identifikations- og diskriminationsgrænserne – uafhængigt af den lingvistiske organisering af forsøgspersonens modersmål. Hvis den kategoriske effekt derimod skyldes sprogspecifikke artikulatoriske strukturer, må den forventes at forekomme ved alle lingvistiske kategorigrænser – og kun dér. Spørgsmålet om hvorvidt den kategoriske effekt skyldes psykoakustiske faktorer eller sprogspecifikke fonetiske kategoriseringer kan ikke besvares helt éntydigt på baggrund af forsøgsresultaterne. Det fremgår dog meget tydeligt, at den lingvistiske organisering af forsøgspersonernes modersmål har en stor indflydelse på identifikationen og diskriminationen af sproglyde fra akustiske kontinua; de to grupper af forsøgspersoner strukturerer de akustiske kontinua forskelligt, men helt i overensstemmelse med de fonetiske kontraster i de respektive sprog. Omvendt viser dele af resultaterne, at sprogspecifikke artikulatoriske strukturer ikke kan forklare alle fænomener, men at sprogbrugere også anvender anden information end de artikulatoriske kommandoer i identifikationen af sproglyde.

1.1 Tegnsystem, kode og medium

Den sproglige kommunikation er en formålsbestemt aktivitet; den er ikke bare lyde eller streger på papir, men også en systematisk organiseret aktivitet motiveret af en intention om at overføre et budskab[1]. Som sådan kan den sproglige kommunikation betragtes ud fra en generel teori om tegn, semiotikken[2], der vedrører studiet af alle aspekter af tegnsystemer til brug for kommunikation.

Ifølge strukturalismen, som den er defineret hos bl.a. Saussure og senere Hjelmslev, er en væsentlig forskel mellem sprogsystemet og andre tegnsystemer, at det sproglige (mindste-)tegn er karakteriseret ved et arbitrært forhold mellem udtryk og indhold[3]. Udtryk og indhold betegnes i Saussures terminologi 'signifié' og 'signifiant', der svarer til en begrebslig størrelse og en akustisk (eller anden fysisk) størrelse:

"The linguistic sign unites, not a thing and a name, but a concept and a sound-image" (Bally & Sechehaye (ed.) 1959:16).

Distinktionen mellem tegnets udtryksside og indholdsside vedrører ikke distinktionen mellem form og substans. Saussure antog, at der er to substanser (lyd og betydning) og én form (Bally & Sechehaye (ed.) 1959). Hos Hjelmslev opdeler denne form den ellers amorfe lyd- og betydningsmasse (der hos Hjelmslev betegnes 'mening', Omkring Sprogteoriens Grundlæggelse 1943, herefter forkortet 'OSG') og så at sige former udtrykssubstansen og indholds-

[1] Dette betyder naturligvis ikke, at den sproglige kommunikation udelukkende er af en propositionel og/eller intentionel karakter. Ifølge Roman Jakobson har den sproglige kommunikation en række funktioner (den emotive, fatiske, poetiske, referentielle, metalingvistiske og den konative funktion), der hver relaterer sig til de konstituerende faktorer (afsender, budskab, kontekst, kontakt, kode og modtager).

[2] Begrebet 'semiotik' (af græsk 'techné sémeiótiké') anvendtes så vidt vides først i antikkens Grækenland af medicinske teoretikere om sygdomssymptomer og -diagnosticering (Morris 1946). De stoiske filosoffer er herefter begyndt at anvende begrebet som reference til en generel tegnteori. I slutningen af det 19. århundrede producerede den amerikanske pragmatiske filosof Charles Sanders Peirce en række skrifter, hvor han definerer semiotik som "the formal doctrine of signs...a sign is something which stands to somebody for something in some respect or capacity" (Hartshorne & Weiss 1931-35: 227-8). Den strukturelle lingvistiks fader, Ferdinand de Saussure, beskriver i begyndelsen af det 20. århundrede semiotikken som en videnskab om tegnenes liv i samfundet og dermed overordnet lingvistikken: "Linguistics is only a part of the general science of semiology" (Bally & Sechehaye (ed.) 1959:16).

[3] Jeg ser her bort fra de få, men ofte omtalte onomatopoietika – de bliver alligevel oftest en del af de arbitrære tegn efter nogen tid (jf. eng. 'pigeon' fra lat. 'pípió').

BEST COPY AVAILABLE

substansen. Hjelmslev pointerede dog i OSG, at heller ikke udtrykssidens form og indholdssidens form er to konforme (ligedannede) plan og at de derfor må beskrives hver for sig. Således definerer strukturalister sprogvidenskaben som en formel videnskab om 'sprogform', hvor genstandsfeltet, sprog, essentielt er et system af former, ikke substans, hvor tegnets afgrænsning, dvs. dets relation til andre tegn, konstituerer form. Hvor form udgør en repræsentation for en lingvistisk enhed, vedrører substans det medium, som den lingvistiske enhed manifesteres i. Manifestationen er altså en funktion mellem form og substans.

Den tidlige strukturalisme, repræsenteret ved Saussure, definerede lingvistikkens objekt som et abstrakt formsystem bestående af tegn og tegnrelationer, men kun tilgængeligt via manifestation i substans. Denne definition af lingvistikkens objekt er senere blevet kritiseret for at lægge for stor vægt på tegnenes relationer i systemet og at negligere den lingvistiske kode[4].

Den lingvistiske kode: Dobbelt artikulation

Det er et definerende karakteristikum ved sprog, at det er uafhængigt af det medium, det manifesteres i (jf. forskellen mellem sprog og tale, language og speech, Sprache og Rede etc.), og at selve den lingvistiske kode udviser en dobbelt struktur. På den ene side har vi betydningsløse elementer med særlige regler for kombinatorik og andre ikke-semantiske egenskaber. På den anden side har vi betydningsbærende kombinationer af de betydningsløse elementer (Lass 1984). Den dobbelte artikulation vedrører altså dannelsen af betydningsbærende enheder, tegn, ved kombinationer af betydningsløse elementer i et forløb.

Sprogets primære medium er lyd

Sprogsystemet er grundlæggende uafhængigt af det medium, det manifesteres i, men skal man fremhæve et medium som det primære eller naturlige for menneskers sprog, er det lyd (Lyons 1981). Jeg vil her antage det synspunkt, at talen er onto- og fylogenetisk primær[5].

[4] På trods af det betydningsfulde fokusskifte fra 'objekter' til 'relationer', der fandt sted i begyndelsen af dette århundrede, er det uheldigt udelukkende at definere et tegns værdi ved dets opposition (eller relation) til andre tegn, dvs. også til deres relationers relationers etc., da denne type definition rammes af en 'infinit regres'. Desuden pointerer Chomsky i 'Aspects' (Chomsky 1965), at det ikke er tilstrækkeligt at beskrive sprogsystemet som et system af elementer og relationer, men at det også må udgøre et system af regler for elementernes kombination.

[5] Med distinktionen mellem talesprog og skriftsprog refererer jeg til det medium, hvori meddelelsen realiseres. Ofte anvendes distinktionen (ikke særlig klart) om normer i tale og skrift, jf. udtrykket "en talesprogsnær tekst" (Basbøll 1977).

Ethvert normalt barn i enhver kultur tilegner sig et talesprog, hvorimod mange aldrig lærer noget skriftsprog. I de tilfælde, hvor personer både tilegner sig et tale- og et skriftsprog, gør de det altid i nævnte rækkefølge (Basbøll 1977). Desuden forholder det sig også sådan, at hvor børn tilsyneladende ubesværet tilegner sig et talesprog, blot de udsættes for sproglig stimulation, så kræver det en god portion motivation og lang tids hårdt arbejde at tilegne sig et skriftsprog (Heger 1992).

Der er sandsynligvis en sammenhæng mellem barnets motoriske udvikling og sprogtilegnelsen: træningen af taleorganerne begynder på et tidspunkt, hvor barnets koordinationsevner og motorik slet ikke har nået et stadium for skriftlig kommunikation. Senere, når barnet når skolealderen og den tid, hvor skriftsproget skal tilegnes, har det allerede tilegnet sig et temmelig nuanceret talesprog. Skriftsproget kommer således til at udgøre en overbygning på det allerede tilegnede (tale-)sprog (Garman 1990).

Alle kendte samfund har ét eller flere talesprog, mens et stort antal af disse ikke har noget skriftsprog (jf. dialekter). Det er ofte tilfældet, at kulturer, der før var analfabetiske, har fået indført et skriftsprog, men ved fremmed indflydelse af en eller anden art.

Endelig er det vigtigt at pointere, at mennesket er biologisk udrustet til at lære at tale et sprog. Hvis et barn vokser op i kontakt med andre mennesker, vil det almindeligvis lære at tale af sig selv, mens man kun med de allerstørste anstrengelser kan lære et højerestående dyr nogle få ord[6].

Fonologiske elementer og taleakter

Distinktionen mellem det formelle sprogsystem og dets manifestation i substans gør sig også gældende i studiet af talen. I studiet af talens udtryksside et det således almindeligt at skelne mellem systemiske former, de fonologiske elementer, og deres applikation i talen, taleakten[7].

[6] De mest kendte forsøg på at lære dyr sproglig kommunikation har været med chimpanserne Vicki, Washoe, Lana, Sarah, Kanzi og Nim Chimpsky. Da opbygningen af chimpansernes taleorganer udelukker muligheden for verbal kommunikation, har man enten anvendt symboltavler eller ASL (American Sign Language). Kun i enkelte tilfælde har man anvendt andre dyr, som f.eks. gorillaer og delfiner. Forsøgene har højst haft en begrænset succes (Steinberg 1993).

[7] Med begrebet 'taleakt' refererer jeg ikke til den pragmatiske teori om sproghandlinger, der hos Austin og Searle betegnes 'Speech Acts' (jf. også Schiffrin 1994; Sperber & Wilson 1986; Mey 1993). Taleakten betegner derimod den konkrete udøvelse af sprogbrug med lyd som medium.

Den strukturalistiske beskrivelse af forholdet mellem taleakten og de fonologiske elementer den antages at repræsentere, udgør traditionelt skillelinjen mellem de to discipliner fonologi og fonetik. Fonologi betegnes ofte som studiet af *former* i talesprog og fonetik som studiet af *substans* i talesprog. Det vil dog, efter min mening, være mere korrekt at sige, at fonologi og fonetik begge vedrører form og substans i talesprog, selvom de varierer i deres primære fokus. Fonologien lægger større vægt på former i talesprog, hvorimod fonetikken først og fremmest vægter talesprogets substans. Netop disse to discipliners afgrænsning, arbejdsfordeling og indbyrdes relationer vil være emnet i det følgende.

Fonologi og fonetik

Distinktionen mellem de to discipliner fonologi og fonetik bygger traditionelt på en antagelse om en ontologisk forskel mellem sprogsystem og sprogbrug. På den ene side siges fonetikken at studere den fysiske eller konkrete virkelighed, taleakten, ved hjælp af naturvidenskabelige (bl.a. akustiske og anatomiske) teknikker. På den anden side siges fonologien at studere den lingvistiske organisering af formelle distinktioner i et konkret sprog.

Hvad der umiddelbart lader til at være en rimelig arbejdsfordeling mellem fonologer og fonetikere afføder ofte diskussioner om talens "egentlige" natur. I lyset af forskellige syn på talens ontologi vægtes områderne mod hinanden og grænserne trækkes hårdt op. En konstatering af, at fonetikken fokuserer på talens universelle fysiske egenskaber, kan alt for let opfattes som en påstand om, at fonetikken vedrører en objektiv fysisk eller konkret virkelighed studeret ved hjælp af videnskabelige metoder, mens fonologien blot forsøger at beskrive en lingvistisk strukturering af denne virkelighed. Omvendt kan samme argument anvendes som grundlag for en påstand om, at fonologien vedrører talens egentlige mentale realitet, hvor fonetikken blot beskriver en fysisk manifestation af denne realitet.

Ovenstående oversimplificerede distinktion mellem fonologi og fonetik er en af de væsentligste årsager til, at relationen mellem disciplinerne let bliver kontroversiel. Effekten er, at grundlæggende definitioner inden for begge områder fokuserer på adskillelse af disciplinerne, og ikke integration af viden.

Det er vigtigt at pointere, at talens universelle karakteristika ikke udelukkende er fonetikkens ejendom, ligesom de sprogspecifikke strukturer ikke alene er

fonologiens. Selvom fonetikere ofte griber til generaliseringer om taleorganer og om auditive og akustiske forhold, er interessen for generaliseringer og kategoriseringer af sprog sjældent mindre hos fonologer. Enhver bestræbelse på at anvende systematiserede begrebssystemer eller notationssystemer (som f.eks. det udbredte IPA-system) i beskrivelsen af forskellige sprog præsupponerer en interesse for universalitet. Omvendt fokuserer en relativt stor del af de fonetiske undersøgelser på sprogspecifikke forhold, f.eks. de danske vokaler[8]. Det vil altså heller ikke være rimeligt at sige, at interessen for universalitet er et definerende karakteristikum for fonetikken, men ikke for fonologien.

Beskrivelsen af den fonetiske analyse som behandling af objektive fysiske eller konkrete forhold er heller ikke entydig. Der er dog ingen tvivl om, at fonetikken i højere grad end fonologien anvender kvantitative metoder; f.eks. er det muligt at røntgenfotografere taleorganerne under udtalen af en vokallyd, og måle muskelaktiviteten samt indlæse de akustiske data i et spektrogram ved hjælp af en spektrograf eller computer. Men disse kvantitative størrelser er meget lidt informative, hvis ikke de kan relateres til sprogbrugerens lingvistiske system. Selv hvis vi er i stand til at redegøre for alle kvantitative forhold, vil vi ikke dermed have afdækket det essentielle ved talen – den er per definition mere og andet end de fysiske forhold.

Omvendt er der også en fare forbundet med en énsidig fokusering på det lingvistiske system; man kan let forledes til at antage, at lingvisten er fri til spekulation om sprogbrugerens mentale organisering af det lingvistiske system. Det er derfor vigtigt, efter min mening, at videnskabelige undersøgelser af sprog – især sprogsystemers mentale realitet – er empirisk funderet, dvs. er baseret på observationer. Med andre ord skal postulater om den mentale organisering af lingvistiske systemer kunne relateres til observerbare fænomener, ligesom kvantitative størrelser (f.eks. artikulatoriske eller akustiske) skal kunne relateres til sprogbrugerens lingvistiske system.

[8] Jf. f.eks. Frøkjær-Jensen 1966 og Petersen 1976.

1.2 Akustisk segmentering og invariante træk

Nogle af de mest centrale spørgsmål i udforskningen af den menneskelige sprogperception er: Hvordan perciperes og fortolkes ytringen på baggrund af taleaktens akustiske signal? Hvilke mekanismer og processer indvirker på perceptionen?

Umiddelbart oplever vi talen som bestående af en serie af diskrete, ikke-overlappende og relativt konstante elementer. Indtil op i 1960'erne var det da også en udbredt antagelse, at man på et tidspunkt – med det rette tekniske udstyr – ville være i stand til at isolere de distinktive elementer, som var de perler på en snor. Man havde store forventninger om at kunne afdække en én-til-én korrespondens mellem det akustiske signal og de fonetiske segmenter[1].

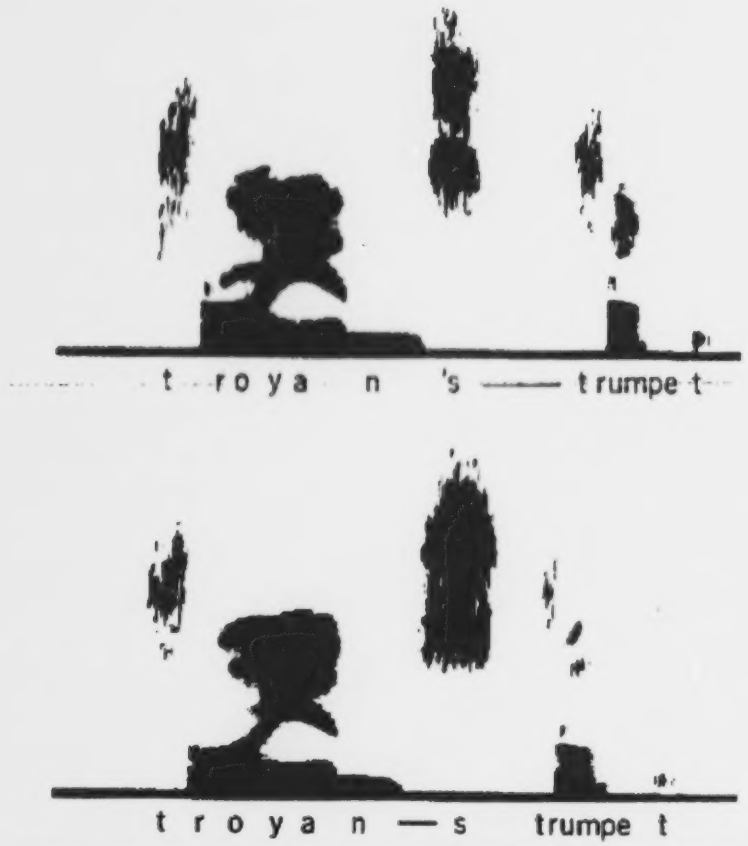
I nogle forsøg baseret på antagelsen om akustisk segmentering optog man en række tydeligt udtalte (engelske) ord: 'pet', 'bat', 'cat' og 'hat' – tydeligt udtalt for at få segmenternes fulde omfang. Derefter ville forsøgslederne isolere [p], [b], [k], [h], [ɛ], [æ] og [t] på magnetbånd. Nu skulle det så i teorien være muligt at kombinere segmenterne i nye mønstre for at konstruere nye ord som f.eks. 'bet' og 'pat' (Peterson et al. 1958).

Forsøgenes resultat bestod udelukkende i en erkendelse af, at det ikke er muligt at isolere fysiske entiteter i talen, der ækvivalerer med fonemer – selv ikke med avanceret teknisk udstyr. De lydlige elementer, der udgør talen, glider over i hinanden og er integreret i hinanden, så det med få undtagelser er meget svært at fastslå, hvor ét segment begynder og et andet slutter (Springer 1979).

Da talen i dens akustiske medium er flygtig og uhåndgribelig, kan det være en stor fordel at anvende en spektrograf, der kan give et visuelt billede af lyden i et frekvens-intensitet-tid-diagram. Figur 1.2.1 er et spektrogram, der repræsenterer taleakterne "Troyans' trumpet" og "Troyan strumpet".

[1] I almindelig dagligdags talesprog kan en taleakt som f.eks. "komme" realiseres fonetisk som [ˈgʰʌmm̩]. Det er min primære hensigt i det følgende at undersøge, hvordan sprogbrugere identificerer rækken af foner i [ˈgʰʌmm̩] fra det akustiske signal. Spørgsmål om, hvordan sprogbrugere – med afsæt i specifikke fonetiske repræsentationer – når frem til (intenderede) ord i en ytring, vil ikke blive behandlet.

BEST COPY AVAILABLE



Figur 1.2.1 Spektrogrammer, der repræsenterer ytringerne "Troyans' trumpet" og "Troyan strumpet". Bemærk det akustiske resultat af forskellen i ordgrænse: frikativen er længere, når den står i initial position, hvilket også gælder for [t]'s aspiration i "trumpet". (Fra Borden & Harris 1984: 195).

Som det tydeligt fremgår, er det ikke muligt i det akustiske signal at afdække isolerbare akustiske segmenter, der ækvivalerer det fonetiske segment. Netop dette forhold viser en fundamental forskel mellem tale og skrift: hvor skriften ofte er segmenteret i alfabetisk ortografi, er talen ikke segmenteret på det akustiske niveau[2].

Fraværet af akustisk segmentering er ikke en ulempe, tværtimod er det en fordel (og måske nødvendighed) ved menneskets tale som verbal

[2] Jeg fokuserer her udelukkende på alfabetisk skrift og ser bort fra bl.a. ikoniske skrifttyper og stavelsesskrift.

BEST COPY AVAILABLE

kommunikation. Vi kan – med et vist besvær – følge en talehastighed på op til 400 ord i minuttet (Orr et al. 1965). Med gennemsnitligt 4-5 fonemer per ord kan vi altså forstå tale med omkring 30 distinktive segmenter pr. sekund. Samtidig ved vi fra den eksperimentelle perceptionspsykologi (Miller 1956), at 30 distinktive lyde langt overskrider ørets temporale kapacitet: f.eks. vokallyde, der overføres ved ca. 20 segmenter pr. sekund, falder sammen i én udifferentierbar tone. Da der selv i almindelig dagligdags talesprog overføres 15-20 distinktive segmenter pr. sekund (dvs. 15-20 fonetiske distinktioner, der adskiller ord, overføres, identificeres og kombineres), er det indlysende, at sprogbrugere umuligt kan kommunikere disse fonetiske distinktioner som isolerede enheder. Endvidere viser en lang række perceptionsforsøg, at kontekstløse sproglyde, f.eks. isolerede vokallyde, kan maskeres af en distraktionslyd op til 250 ms efter præsentationen (Massaro 1970, 1971, 1972, 1975). I (danske) forsøg, hvor forsøgsdeltagerne selv må tilpasse fuldvokalers længde, viser det sig, at de foretrækker en varighed over 200 ms for en isoleret langvokal (Petersen 1976). Kort sagt kræver identifikation af isolerede sproglyde så lang tids processering i forhold til talehastigheden, at akustisk segmentering ville være alt for tidskrævende i normale kommunikationssituationer.

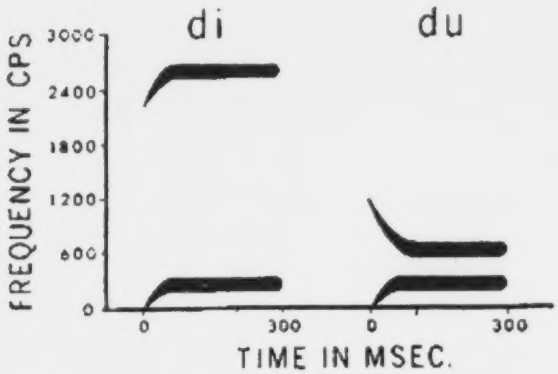
Akustisk integration

Man kunne have forestillet sig, at det var muligt at finde konstante akustiske træk til identifikation af fonemet, selvom fonemet ikke er en selvstændig, isolerbar akustisk enhed. Dette er dog ikke tilfældet. Undersøgelser foretaget med Vocoder talesynthesizere[3] giver endnu et fingerpeg om, hvorfor det generelt er umuligt at isolere fonetiske segmenter i talen.

Figur 1.2.2 viser spektrogrammer, der repræsenterer de to konsonant-vokal-stavelser (herefter CV-stavelser) [di] og [du]. De eneste variabler, der behandles i synthesizeren, er de to formantfrekvenser samt grundtonen, der bliver holdt konstant ved alle stimuli. De mørke bånd i spektrogrammerne repræsenterer henholdsvis første og anden formantfrekvens (herefter F1 og

[3] Den første Vocoder-synthesizer blev udviklet af H. Dudley og hans kollegaer ved Bell Telephone Laboratorierne i slutningen af 1930'erne. Den blev oprindeligt udviklet med det formål at reducere båndbredden i f.eks. oversøiske telefonkabler, men blev aldrig nogen rigtig succes på det område. Derimod satte Vocoderen gang i udviklingen af talesynthesizere og havde derved stor indflydelse på den senere udforskning af talesyntese og sprogperception (Lieberman & Blumstein 1988).

F2) som funktioner af tid. De to første formanters relative frekvens er tilstrækkelig information til identifikation af vokalen[4].



Figur 1.2.2 Spektrogrammer over syntetiske transitionsmønstre, der repræsenterer CV-stavelserne [di] og [du]. Spektrogrammerne er håndmalede og kan aflæses af en Vocoder talesynthesizer, der konverterer dem til lyd. F1- og F2-information er medtaget, mens grundtonen ikke er repræsenteret i spektrogrammerne. Bemærk at de to F2-transitioner peger i forskellig retning. (Fra Liberman, Cooper, Shankweiler & Studdert-Kennedy 1967: 435).

F1 og F2's frekvens i naturlig tale bestemmes af taleorganernes – især tungens – position under udtalen; tungens relative højde determinerer F1's frekvens, hvorimod F2's frekvens determineres af tungens relative horisontale position, dvs. hvorvidt der er tale om en fortunge- eller en bagtungevokal. F.eks. er tungen hævet mod ganen ved udtalen af både [i] og [u], hvilket betyder, at deres F1 skal findes i nogenlunde samme frekvensområde. Derimod er der relativ stor forskel i tungens position ved udtalen af henholdsvis [i] og [u]: [i] udtales med hævet fortunge, hvilket resulterer i en F2 nær 3000 Hz, mens [u] udtales med hævet bagtunge, hvilket resulterer i en F2 omkring 700 Hz.

I begyndelsen af hver CV-stavelse bevæger formanterne sig meget hurtigt gennem en række frekvenser. Ved [di] er der tale om en stigning fra omkring 2200 Hz til 2600 Hz, hvor det i [du] drejer sig om et fald fra omkring 1200 til 700 Hz. De hurtige frekvensskift, der hver varer ca. 50 millisekunder, kaldes formanttransitioner. Det er netop disse formanttransitioner, som lytteren

[4] Det skal dog tilføjes, at vokallyde, der udelukkende består af to formanter, ikke lyder særlig naturligt; de lyder snarere som dut-lyde end vokallyde.

BEST COPY AVAILABLE

identificerer, når han/hun "hører" 'd' (i modsætning til 'b' eller 'g') i stavelserne. Det er her værd at bemærke, at F2-transitionen i [di] og [du] er rettet fra samme frekvensområde, nemlig 1800 Hz. Dette skyldes, at F2-transitionens locus[5] er determineret af konsonantens artikulationssted, og artikulationsstedet lige før eksplosionen i [d] er omtrent det samme før alle vokaler. Straks efter eksplosionen i [d] begynder taleorganerne at glide mod vokalens artikulationsposition. Det er lige netop denne glidende overgang fra én artikulationsposition til en anden, der skaber formanttransitionerne[6]. Siden artikulationspositionerne for [i] og [u] er væsentlig forskellige, er F2-transitionerne før vokalerne også forskellige (Segalowitz & Gruber 1977).

Det er en vigtig pointe, at det akustiske træk, der determinerer perceptionen af [d], F2-transitionen, er væsentlig forskellig foran [i] og [u]. Konsonanten [d] siges at være integreret i den efterfølgende vokal[7]: det er f.eks. ikke muligt at producere transitionen uafhængigt af en vokal; transitionen eksisterer kun i kraft af relationen til vokalen; der er ikke et specifikt akustisk segment, som svarer til det perciperede /d/-fonem. Den akustiske realisation af f.eks. /d/ i én stavelse er altså noget anderledes end i en anden stavelse. Det er derfor heller ikke muligt at sætte [d]'s formanttransitioner i [di] sammen med vokalen [u] for derved at konstruere stavelsen [du].

Denne specifikke egenskab ved integration – at der ikke eksisterer en korrespondens mellem det akustiske signal og individuelle fonemer – regnes for at være det mest essentielle karakteristikum ved talen (Lieberman & Bloomstein 1988). Liberman og hans kollegaer i Haskins-gruppen er nok dem der mest eksplicit har redegjort for dette synspunkt:

"Acoustic cues for successive phonemes are intermixed in the sound stream to such an extent that definable segments of sound do not correspond to segments at the phoneme level.

[5] Forholdet mellem F2-transitioners loci – eller "begyndelsesfrekvenser" – og artikulationssted er blevet studeret indgående bl.a. i Delattre et al. 1955.

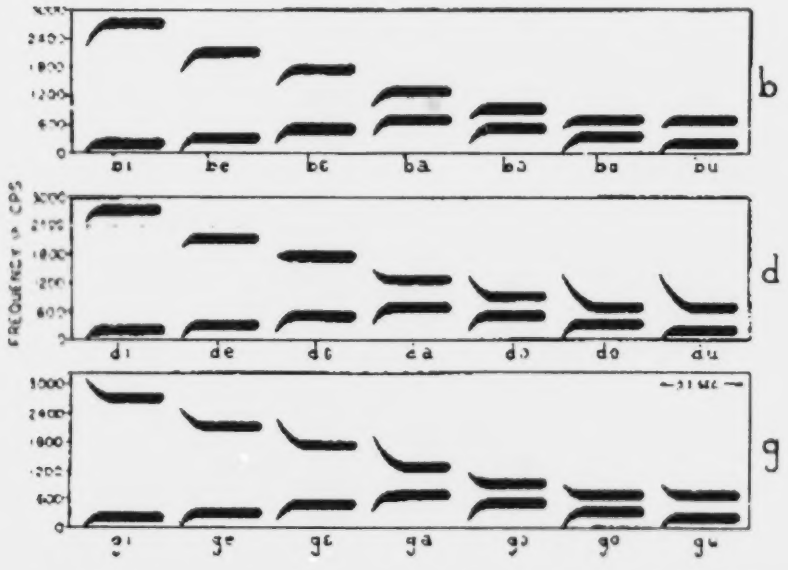
[6] Formanttransitionerne forårsages ikke af vokalens stemthed; selv i taleakter uden stemthed, f.eks. når der hviskes, er det muligt at følge formanttransitionerne.

[7] Det særlige forhold, at konsonanters formanttransitioner er en integreret del af vokalen, kaldes i den engelsk-amerikanske litteratur for "encodedness" (se f.eks. Liberman et al. 1967; Segalowitz & Gruber 1977; Lieberman & Bloomstein 1988; Borden & Harris 1984, etc.). Jeg har her valgt betegnelsen "integration", da der snarere er tale om effekten af artikulatorisk inerti end en egentlig intentionel kodning. Det kan undre mig, at Liberman et al. explicit har valgt at anvende en terminologi, der stammer fra et helt andet videnskabsparadigme, nemlig kryptografien, og derved antyder, at der i talen skulle eksistere en arbitrær relation mellem en original og en kodet form (jf. Liberman et al. 1967).

Moreover, the same phoneme is most commonly represented in different phonemic environments by sounds that are vastly different. There is, in short, a marked lack of correspondence between sound and perceived phoneme" (Liberman, Cooper, Shankweiler & Studdert-Kennedy 1967: 432).

Varians

Som det fremgik af det ovenstående, er segmenterne i det akustiske signal integreret i hinanden med den konsekvens, at det ikke er muligt at afdække en én-til-én korrespondens mellem det akustiske signal og de fonetiske segmenter. Der er altså kun en indirekte relation mellem den fonetiske repræsentation og det akustiske signal, som den udspringer af, og denne relation er yderst kompliceret. På den ene side kan forskellige akustiske træk producere samme fonetiske kontrast, men på den anden side kan flere fonetiske kontraster også manifesteres ved samme akustiske træk (Fitch et al. 1980).



Figur 1.2.3 Spektrogrammer over de syntetiske transitionsmønstre for [b], [d] og [g] i forskellige CV-stavelser. Bemærk den kontekstbetingede variation i F2-transitioner . (Fra Delattre, Liberman & Cooper 1955).

BEST COPY AVAILABLE

Mange-til-én korrespondens mellem det akustiske signal og det fonetiske segment skyldes oftest kontekstbetinget variation i de akustiske træk. Et eksempel på variation i det akustiske signal er figur 1.2.3, der viser en serie spektrogrammer af [b], [d] og [g] i forskellige CV-stavelser. Det er især værd at bemærke, at F2-transitionens retning i "samme" konsonant varierer afhængigt af den efterfølgende vokal.

En-til-mange korrespondens mellem det akustiske signal og det fonetiske segment er ligeledes en konsekvens af de akustiske segmenters integration. I et bemærkelsesværdigt forsøg udført af Carol D. Schatz (Schatz 1954) havde hun splejset eksplosionen fra naturligt producerede CV-stavelser sammen med vokalerne [i], [a] og [u]. Ved præsentation for en række forsøgspersoner blev samme naturligt produceret eksplosion ved 1440 Hz identificeret som henholdsvis [p] og [k] afhængigt af, om den efterfølgende vokal var [i] eller [a][8]. Liberman, Delattre og Cooper (1952) havde tidligere nået til samme resultat med syntetiske stimuli, hvor eksplosionen ved 1440 Hz blev identificeret som [k] foran [a], men som [p] både foran [i] og [u].

Akustisk invarians

En af konsekvenserne af de akustiske segmenters integration er altså, at det akustiske signal ikke kan segmenteres i størrelser, der svarer til fonemer. Spørgsmålet må herefter være, om der eksisterer invariante akustiske størrelser, som sprogbrugere kan anvende i perceptionen af taleakter realiseret i et akustisk signal.

Det er min hensigt i det følgende at opsummere nogle af de vigtigste og mest relevante resultater af forsøg på at afdække invariante træk i det akustiske signal. Resultaterne vil blive præsenteret i afsnit, der dækker forskellige områder som f.eks. 'locus', F2, F3 og friktionsstøj (dækker her eksplosionsstøj, aspirationsstøj og friktionsstøj i traditionel forstand).

Resultaterne i det følgende stammer både fra forsøg med syntetiske stimuli produceret ved hjælp af talesynthesizere, naturlig tale samt spektrografiske analyser.

[8] I Liberman, Delattre og Coopers (1952) oprindelige forsøg var de syntetiske stimuli produceret ved hjælp af en talesynthesizer. Carol D. Schatz anvendte derimod naturlig tale, som blev redigeret i en båndmaskine.

En lang række undersøgelser af de akustiske træk, som gør det muligt for sprogbrugere at identificere stavelser og ord, bygger på kontrollerede eksperimentelle metodologier. Typisk varieres et akustisk træk i stavelser eller ord, der i anden henseende er invariante. Disse stimuli præsenteres derefter under nær optimale forhold for en række forsøgspersoner, som bliver bedt om at identificere de præsenterede stimuli. På den måde er det muligt at afgøre et akustisk træks relevans for fonetisk perception (Bond 1981).

Locus[9]: en akustisk invarians?

Antagelsen om loci (eller formanters hypotetiske begyndelsesfrekvenser) som mulige invariante akustiske træk tager afsæt i det problem, som vi blev præsenteret for i figur 1.2.3, nemlig at formanter i "samme" konsonant varierer med den efterfølgende vokal.

Ændringer i formanttransitionernes grafiske repræsentation skyldes, at taleorganerne i løbet af taleakten glider fra en position mod en anden. Med andre ord afspejler formanttransitionerne ændringer i mundhulrummets størrelse og form forårsaget af taleorganernes bevægelse. Da konsonanternes artikulationssted er relativt konstant, er det en nærliggende tanke, at man må kunne beskrive deres locus – eller "begyndelsesfrekvenser".

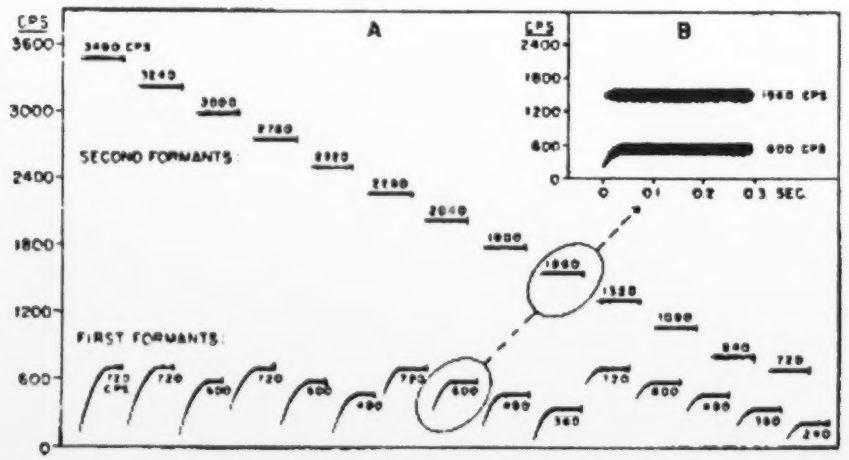
Hvis man i figur 1.2.3 ekstrapolerer de forskellige F2-transitioner, som karakteriserer den enkelte konsonant, tilbage i tid, ser vi, at de alle peger mod samme frekvensområde. Ud fra en antagelse om, at transitionernes første del mangler, kan man altså reducere de mange forskellige F2-transitioner til ét artikulatorisk og akustisk locus.

I figur 1.2.3 er [d] tydeligvis det bedste eksempel med en konstant F2-frekvens på 1800 Hz. F2-transitionerne i [b] er derimod alle stigende, hvilket betyder, at [b]'s locus skal findes i et lavere frekvensområde end den laveste F2-frekvens (dvs. < 800 Hz). Det mest komplicerede eksempel er [g], hvor det ikke er muligt at reducere F2-transitionerne til ét locus. Man må her antage, at der er et højfrekvens-locus på 3000 Hz eller derover for

[9] Da locusbegrebet oftest anvendes om artikulationssted og (det akustiske resultat) F2, kan man let forledes til at tro, at locusbegrebet er begrænset til disse områder. Med 'locus' refereres der ganske enkelt til det frekvensområde, som man antager, en transition udspringer af, og som transitionens grafiske repræsentation peger mod – uanset transitionens årsag (jf. Delattre et al. 1955).

fortungevokalerne [i], [e] og [ɛ], og et lavfrekvens-locus for bagtungevokalerne [ɔ], [o] og [u] omkring 1200 Hz.

I en række forsøg med stimuliene [b], [d] og [g] forsøgte Delattre, Liberman og Cooper i 1955 at finde eksperimentel evidens for teorien om loci som invariante akustiske træk. Hensigten med undersøgelserne var at fastslå (1) om loci er kontekstuafhængige; (2) om locusbegrebet er relevant i forbindelse med F1-transitioner; (3) om loci udgør transitioners begyndelsesfrekvenser eller om loci blot er frekvensområder, transitionerne er rettet mod.



Figur 1.2.4 Skema over stimuli, der blev anvendt i forsøget på at finde [b], [d] og [g]'s F2-loci. **A**: 13 konstante F2-frekvenser og 5 F1-frekvenser med transition, der kombineret udgjorde 65 stimuli. **B**: Et eksempel på en typisk teststimulus bestående af de to formanter, der er markeret i 'A'. (Fra Delattre, Liberman & Cooper 1955).

I første omgang anvendte Delattre et al. en teknik, hvor de præsenterede stimuli – i alt 65 – bestående af (5 · 13) F1-F2-kombinationer. Som det fremgår af figur 1.2.4 var alle F2 uden transitioner, hvilket var nødvendigt for at få F2 til at udpege (i bogstavelig forstand) locusfrekvenserne. Når F2 er uden transition, er det desuden tvingende nødvendigt, at F1 har transition, hvis en del af en stimulus overhovedet skal kunne identificeres som en klusil; ellers vil den blive identificeret som en vokalisk lyd.

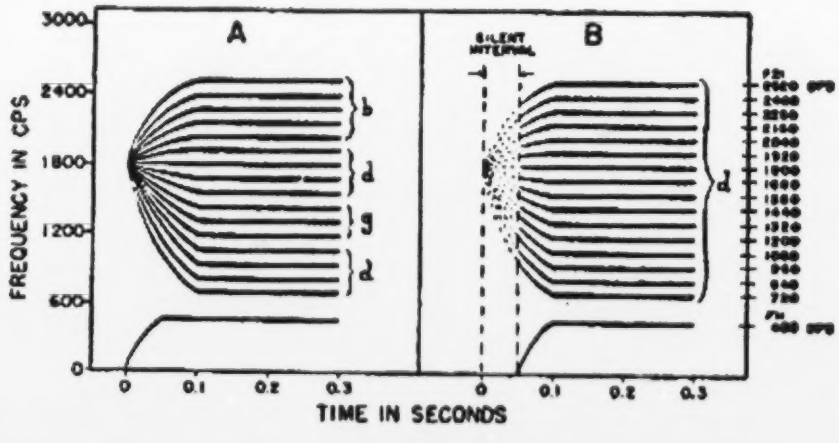
Resultatet af første forsøg svarer meget godt til de ekstrapolerede F2-transitioner i figur 1.2.3: det bedste eksempel på [g] har en F2-frekvens på

BEST COPY AVAILABLE

3000 Hz, det bedste [d] på 1800 Hz og [b] på 720 Hz. Andre F2-frekvenser gav en mere eller mindre vellignende klusillyd afhængigt af afstanden til en af de tre ovennævnte F2-loci. Ved en F2 på 1320 Hz var lyden et udifferentierbart [b], [d] eller [g]. Med hensyn til F1-transitionens betydning for de tre loci er konklusionen entydig:

".. it appears that the second-formant loci are independent of the changes in vowel color that are produced by varying the position of the first formant" (Delattre et al. 1955:92).

I en anden undersøgelse forsøger Delattre et al. at fastslå, om F2-transitionen begynder ved dens locus, hvorefter den glider over i en steady-state vokal, eller om F2-transitionen blot er rettet mod dens locus. Som det fremgår af figur 1.2.5 anvendte Delattre et al. et af resultaterne fra den føromtalte undersøgelse, nemlig frekvensen for [d]'s locus, som det blev fastslået er på 1800 Hz. Rækken af stimuli varierer med hensyn til F2's bevægelse fra 1800 Hz til en steady-state vokal, hvorimod F1 er holdt konstant med et onset, der er synkroniseret med F2's onset.



Figur 1.2.5 Skema over stimuli og deres identifikation – med og uden pause mellem locus og transitionernes onset. **A**: F2-transitioner, der begynder ved [d]'s locus og derefter glider mod forskellige steady-state vokaler. F1 er konstant for alle stimuli. **B**: Samme stimuli som i 'A', blot med en 50 ms pause mellem locus og F2-transitionens onset. Bemærk at der ikke er nogen pause i F1, men at F1 er forskubbet, så dets onset er synkroniseret med F2's onset. (Fra Delattre, Liberman & Cooper 1955).

Resultatet af identifikationstesten er bemærkelsesværdigt: på trods af at F2-transitionen i alle stimuli begynder i [d]'s locus, blev ikke alle stimuli

BEST COPY AVAILABLE

identificeret som [d] (jf. figur 1.2.5 A.). Derimod blev stimuli med en F2-frekvens mellem 2040 og 2520 Hz identificeret som [b], mellem 1560 og 1920 Hz som [d], mellem 1200 og 1440 Hz som [g], og under 1200 Hz igen som [d].

Da Delattre et al. derefter indsatte en pause mellem locus og transitionens onset ved at slette de første 50 ms (jf. figur 1.2.5 B.), blev alle stimuli identificeret som et mere eller mindre vellignende [d]. Forsøg med stimuli, der tog udgangspunkt i henholdsvis [b] og [g]'s locus, gav lignende resultater:

"It is reasonably clear in these cases, as it was with *d*, that the transition cannot start at the locus and go all the way to the steady state" (Delattre et al. 1955: 94).

Om transitionen begynder ved locus eller blot er rettet mod locus er alt-afgørende for relevansen af locus som invariant akustisk træk. Der er ingen tvivl om, at locus kan defineres akustisk, altså som et bestemt frekvens-område, og artikulatorisk som et bestemt artikulationssted. Men da det er nødvendigt at slette den første del af transitionen, så den ikke begynder i locus, kan locus umuligt udgøre et relevant, invariant træk for identifikation af konsonanten. Locus er simpelthen ikke en (relevant) del af det akustiske signal.

Derimod er der ingen tvivl om, at locus er en reel artikulatorisk konstant. Det artikulatoriske lukke er konstant for alle [d] i CV-stavelser relativt uafhængigt af den efterfølgende vokal. Ifølge beregninger foretaget af Stevens & House (1956) er mundhulrummets resonansfrekvens ved lukket på omkring 1800 Hz. Men da der ikke produceres lyd, før taleorganerne er begyndt at åbne, indgår locus ikke direkte i det akustiske signal. Selvom locus umiddelbart er et mere invariant træk for fonemet end transitionerne selv, er dets invarians afledt og nærmere relateret til artikulatoriske forhold end akustiske.

Akustisk invarians i F2-transitioner

I det foregående har jeg beskrevet den varians, der kendetegner klusilerne. Et af de væsentligste akustiske træk for perceptionen af klusiler er F2-transitionerne, der signalerer artikulationssted. Manglen på invarians i F2-transitioner, som f.eks. i [d], er karakteristisk for såvel aspirerede som uaspirerede klusiler samt nasalerne (Liberman et al. 1954). Det er i disse konsonanter, at "the acoustic-phonetic noninvariance problem" kommer tydeligst til udtryk (Klatt 1979).

Den største grad af diskontinuitet skal nok findes i de velære lukkelyde [k], [g] og [ŋ]. Ikke alene har de varierende F2-transitioner afhængigt af den efterfølgende vokal, men også varierende loci, der er betinget af konteksten. F.eks. er der et betydeligt skift i konsonantens locus-frekvens mellem CV-stavelser med henholdsvis rundede og urundede vokaler (Liberman et al. 1967). Disse forhold er netop med til at skabe en kompleks relation mellem det akustiske signal og de fonetiske segmenter.

Hos sonoranterne [r], [l], [w] og [j] ser vi ikke samme kompleksitet mellem akustisk signal og fonetisk segment, da F2-transitionen i disse konsonanter har deres begyndelse ved et relativt invariant locus[10], hvilket ikke er tilfældet med klusiler og nasaler. Derimod er F2-transitionerne i sig selv ikke spor konstante – også her er variationen betinget af den efterfølgende vokal (O'Connor et al. 1957).

Akustisk invarians i F3-transitioner

Er det så muligt at finde invariante træk for perception i F3-transitionerne? Vi må starte med at konstatere, at der ikke er blevet forsket særlig meget i F3-transitioner og deres betydning for perceptionen. Grunden til dette er, at F3 øjensynligt ikke har den store betydning for andet end perceptionen af nasale konsonanter (Lisker et al. 1957), og selv for de nasale konsonanter er der ikke noget, der tyder på, at F3-transitionen udgør et invariant akustisk træk (Harris et al. 1958; O'Connor et al. 1957; Liberman et al. 1967).

Faktisk synes "the acoustic-phonetic noninvariance problem" at blive yderligere kompliceret, da F3-transitionens vigtighed er relativ i forhold til den fonetiske kontekst – især F2-transitionerne. I eksempler som [di]-[du], jf. figur 1.2.2, fandt Liberman et al. (1967), at en passende F3-transition forbedrer muligheden for en korrekt identifikation af [d] i [di], men slet ikke i [du]. Det er sandsynligt, at F3-transitionen indvirker på perceptionen i bestemte formant-konstellationer, men det er samtidig muligt at fastslå, at F3 i endnu mindre grad end F1 og F2 kandiderer til rollen som invariant akustisk træk.

[10] Dette er i høj grad tilfældet for [w], men mindre for [r], [l] og [j], hvilket skyldes artikulatoriske forhold.

Invarians i friktionsstøj[11]

Støj, der dannes som friktionsstøj i taleakten, hvad enten den findes som længere passager i frikativer eller som kortvarig eksplosionsstøj i klusiler, er et andet akustisk træk for identifikation af konsonanterne – især deres artikulationssted.

På den ene side har vi en række tilfælde, hvor friktionsstøjen er relativt lang, typisk i frikativer, som almindeligvis ikke er underlagt den store kontekstbetingede variation[12]. Siden det akustiske træk, der er forbundet med friktionsstøjen, er altoverskyggende i perceptionen af [s] og [š], må vi konstatere, at disse frikativer udviser en lav grad af restrukturering i det akustiske signal. Friktionen udgør med andre ord en undtagelse, da den ikke i samme grad er integreret i de omkringliggende akustiske segmenter, som vi tidligere har set det med mange andre akustiske træk (Liberman et al. 1967).

På den anden side har vi også friktionsstøj i klusilerne, hvor en kortvarig eksplosionsstøj udgør det akustiske resultat af artikulationen. Klusilernes eksplosionsstøj er dog mindst lige så restruktureret og diskontinuert et træk som klusilernes formanttransitioner.

I forsøg udført af Liberman et al. (1952) har man forsøgt at producere de bedst mulige klusiler [k] og [g] ved hjælp af syntetiske spektrogrammer. Her viser det sig, at eksplosionsstøjens frekvens er direkte relateret til den fonetiske kontekst – i dette tilfælde den efterfølgende vokal . Faktisk varierer eksplosionsstøjens frekvens over så stort et område, at det overlapper med det frekvensområde, som producerer det bedste [b] og [p]. Resultatet af dette er, som tidligere beskrevet i afsnittet "Varians", at der opstår en én-til-mange korrespondens mellem det akustiske signal og det fonetiske segment, f.eks. identificeres en eksplosionsstøj ved 1440 Hz som henholdsvis [p] foran [i] og [k] foran [a].

Endelig skal det pointeres, at eksplosionens vigtighed er relativ og varierer for samme klusil i forskellige kontekster, hvilket også gælder for F2-transitionen. F.eks. er det nødvendigt for [g] og [k], der er markeret med en kraftig F2-transition før [æ], at deres eksplosionsstøj falder inden for et nøje

[11] Friktionsstøjens relevante fysiske variabel er et (i spektrogrammer lodret) intensitetsbånds frekvens.
[12] Dog assimileres rundingen fra en efterfølgende vokal; sammenlign f.eks. [s] i sile-syle og stirre-styre.

afgrænset frekvensområde, hvis de skal kunne identificeres foran [o][13]. Med andre ord kan samme konsonant i forskellige kontekster afhænge af så forskellige akustiske træk som friktionsstøj og formanttransitioner[14].

[13] Disse resultater stammer fra forsøg med eksplosionsstøj og transitioner som akustiske træk for identifikation af klusiler (Liberman et al. 1952, 1954).
[14] Se Repp (1982) for en gennemgang af forskellige typer 'trading relations'.

1.3 Teorier om sprogperception

Forudsat at der ikke eksisterer diskrete, invariante træk for de distinktive fonemer, rejser der sig nogle helt centrale spørgsmål: (1) hvordan forholder de fonetiske segmenter sig til det akustiske signal? og (2) ved hvilke processer restrukturerer lytteren de fonetiske segmenter fra det akustiske signal?

To generelle tendenser i dannelsen af teorier om sprogperceptionen har været dominerende i den videnskabelige debat. I den ene gruppe af teorier (her repræsenteret ved bl.a. signal-detektor-teorien, jf. side 43ff) beskrives lytteren som relativt passiv og selve perceptionsprocessen som primært sensorisk. Det akustiske signal modtages, filtreres og sammenstilles så direkte med de akustiske-fonetiske træk på det pågældende sprog.

Den anden gruppe af teorier (synteseteorierne, her repræsenteret ved Motorteorien og Analysis-by-Synthesis-teorien) anser lytteren for at være mere aktiv og hævder desuden, at perceptionsprocessen indeholder visse aspekter af produktionsprocessen; det akustiske signal modtages og analyseres for fonetiske egenskaber i henhold til, hvordan de enkelte lyde produceres.

Synteseteorier

Der er to generelt accepterede teorier om fonetisk sprogperception, som fokuserer på relationen til sprogproduktion[1]: Motorteorien, repræsenteret ved Alvin M. Liberman og en række af hans kollegaer ved Haskins-Laboratorierne i New Haven, Connecticut, og Analysis-by-Synthesis-teorien repræsenteret ved Kenneth N. Stevens og Morris Halle ved Massachusetts Institute of Technology.

Begge teorier udspringer af de store forventninger om at kunne afdække en én-til-én korrespondens mellem det akustiske signal og fonemer, man havde i den første halvdel af dette århundrede. Men det er svært – hvis ikke umuligt – at finde konstante akustiske træk til identifikation af fonemer. Det var derfor nødvendigt at udvikle en model, der omfatter fonemer i de

[1] En tredje, men ikke udbredt teori bygger på Gibson-paradigmet (Gibson 1966), hvor alle former for perception sammenholdes med stimulationens oprindelse. Hos Gibson anses perceptionen af sprog ikke som noget grundlæggende andet end perception af auditive (og visuelle) signaler, men taleorganernes specielle natur fremhæves og sprogperceptionen regnes derfor som speciel, fordi sprogets oprindelse er speciel (jf. også C.A. Fowlers 'Direct-Realist' perspektiv, Fowler 1986). Denne teori vil ikke blive yderligere uddybet her, da den har en perifer status i den aktuelle debat om sprogperception.

perceptuelle processer, men som ikke kræver, at fonemerne identificeres ud fra diskrete, invariante træk i det akustiske signal[2].

I en tidlig formulering af en perceptionsteori baseret på syntese, antog man, at talegenkendelse var baseret på, at det akustiske signal kunne matches med en idealiseret form, en skabelon[3]. Man antog, at modtageren, efter at have perciperet det akustiske signal, begyndte at generere eller syntetisere en psykoakustisk skabelon, som så kunne sammenholdes med det perciperede.

Med den tidlige synteseteori løb man hurtigt ind i et centralt problem: der er mere end én taler i verden. Sammenstillingen af akustiske træk med fonetiske segmenter – også kaldet "the acoustic-phonetic mapping" (Clark & Clark 1977: 205) – varierer fra taler til taler, og endda fra taleakt til taleakt hos samme taler. Enhver form for variation, hvad enten det drejer sig om inter- eller intraindividuel variation, fører til en ny akustisk-fonetisk sammenstilling. Da variationen er potentielt infinit, ville det være nødvendigt med et infinit antal af disse skabeloner, hvis en positiv identifikation af det perciperede skulle være mulig (Harley 1995).

Netop spørgsmålet om, hvordan det perciperede signal matches med en skabelon og problemer vedrørende skabelonens karakter, har ført til en række forskellige løsninger og modeller inden for synteseparadigmet. Fælles for teorierne er dog, at de bygger på en antagelse om, at lytteren er aktiv deltager i en proces, hvor det akustiske signal matches med en idealiseret form; om den idealiserede form så er af en akustisk eller en motorisk karakter er netop det skelsættende.

Motorteorien

En af de mest kendte og omdiskuterede varianter af synteseteorien er Motorteorien. Den oprindelige Motorteori tog, som de andre synteseteorier, udgangspunkt i en formodning om, at der eksisterer en direkte korrespondens mellem det akustiske signal og de fonetiske segmenter. Undersøgelser viste

[2] På trods af, at tanken om, at fonemer kan identificeres ved syntese af invariante akustiske træk (altså en aktiv perceptionsproces) ikke regnes for plausibel i moderne psykolingvistik og perceptionsteori, eksisterer der flere fremtrædende varianter af teorien. Den mest kendte og udbyggede model er Klatts LAFS model – Lexical Access From Spectra (Klatt 1989).

[3] Jeg har valgt at anvende betegnelsen 'skabelon' for 'template', som almindeligvis anvendes i den engelsksprogede litteratur. Efter min opfattelse associerer 'skabelon' med en idealiseret fysisk form eller blueprint, hvilket netop svarer til de tidlige beskrivelser af 'template matching'.

dog snart, at det ikke er tilfældet; det var ikke muligt at give en entydig akustisk beskrivelse af alle fonemer. Fraværet af akustisk invarians – det såkaldte 'acoustic-phonetic noninvariance problem' – udgjorde et grundlæggende problem for teorier, der antog, at talen består af sekvenser af fonetiske realisationer. Med Motorteorien indledtes i 1950'erne og 1960'erne et forsøg på at fastholde fonemet som en perceptuel enhed.

Det centrale punkt i Motorteorien er, at den fokuserer på, at lytteren på andre tidspunkter er taleren og at lytteren anvender sin viden om artikulatoriske forhold i afkodningen af det akustiske signal[4]. Med andre ord bygger perceptionsmodellen på den hypotese, at forholdet mellem det akustiske signal og de fonetiske segmenter medieres af mekanismer, der indgår i produktionen af tale, artikulationen, hvorfor teorien generelt er kendt som 'the Motor Theory of Speech Perception':

"The listener uses the inconstant sound as a basis for finding his way back to the articulatory gestures that produces it and thence, as it were, to the speaker's intent" (Liberman, Cooper, Shankweiler & Studdert-Kennedy 1967: 453).

Da akustisk kodning fortrinsvis forårsages af koartikulation (jf. kap. 2.1), er det da også nærliggende, at viden om artikulation og taleorganernes dynamik har en væsentlig funktion i afkodningen af det akustiske signal.

Motorteoriens to hypoteser

Motorteorien har fra begyndelsen af 1950'erne til i dag gennemgået en stadig udvikling og ændret selv centrale postulater i samspil med empiriske undersøgelser og fagfolks kritik. Teoriens udvikling er selv af kritikere blevet karakteriseret som kontinuerlig indtil 1985, hvor teoriens grundlæggende princip (the Motor Command Principle) blev udsat for en gennemgående revision. Jeg mener derfor, at det er rimeligt at fremlægge Motorteoriens udvikling i to faser, inddelt efter omfanget af teoriens postulater, som jeg vil referere til som henholdsvis 'stærk motor hypotese' og 'svag motor hypotese' (Clark & Clark 1977; Liberman & Mattingly 1985; Mattingly & Studdert-Kennedy 1991).

[4] Der er nok ingen tvivl om, at vi anvender vores viden om artikulationsforhold, når vi lytter. Det er en anerkendt og dokumenteret antagelse, at vi også trækker på det visuelle i tolkningen af det akustiske signal. De visuelle signaler, som f.eks. læberunding ved vokaler, er dog hverken nødvendige eller tilstrækkelige, men blot en hjælp i fortolkningen af almindelig dagligdags tale, og de vil derfor ikke blive behandlet nærmere her.

Et tidligt udkast til Motorteorien antog en stærk hypotese, der bygger på en i dag absurd antagelse, nemlig at lytteren producerer en slags "indre monolog", der så kan matches med det perciperede signal for identifikation (Crystal 1995). En lang række undersøgelser og forsøg taler kraftigt imod en antagelse som denne.

For det første er der intet der tyder på, at sprogbrugere producerer en eller anden form for subvokal monolog, mens man taler til dem. Der er simpelthen ingen former for aktivitet i de muskler, der almindeligvis indgår i artikulationen.

For det andet, hvis mekanismer der indgår i taleproduktion var en nødvendig betingelse for en perceptionsproces, ville personer, som af neurologiske årsager ikke kan tale, heller ikke være i stand til at forstå tale. Bl.a. Lenneberg (1962) har rapporteret en række tilfælde med personer uden tale, der ikke har nogen problemer overhovedet med at forstå tale. Og mere generelt implicerer Motorteoriens stærke hypotese, at sprogbrugere ikke kan forstå mere end de kan producere[5]. Det forholder sig snarere lige omvendt. Vi kan alle almindeligvis forstå sprogbrugere, som har en anden udtale end vores egen. F.eks. har personer der stammer, mandlige sprogbrugere, børn, dialekttalende, personer med udenlandsk accent, etc. ikke særlige forståelsesproblemer, når de kommunikerer med voksne kvindelige sprogbrugere, der ikke stammer eller har en særlig accent.

Den oprindelige Motorteori, som beskrevet i Cooper et al. (1952), Delattre et al. (1955) og Liberman et al. (1959), og senere uddybet i Liberman et al. (1967), udvikledes således til en mere generel perceptionsteori med en lineær model over, hvordan lyttere kan identificere segmentale fonemer uden reference til invariante akustiske træk: På trods af en ringe korrespondens mellem det akustiske signal og de fonetiske segmenter, indebærer produktionen af fonemer en reference til et sæt invariante motoriske kommandoer; perceptionen af fonemer kan således også medieres af en reference til de samme invariante motoriske kommandoer.

[5] Dette gælder i hvert fald ikke personer, der er stumme. Mere generelt virker det i øvrigt usandsynligt, at sprogbrugere ikke kan forstå mere end de kan producere. I forbindelse med sprogtilegnelse forholder det sig snarere lige omvendt, hvadenten det drejer sig om modersmålstilegnelse, andet- eller fremmedsprogstilegnelse. Vi kan almindeligvis skelne mellem meget mere end vi kan identificere, og vi har også lettere ved at identificere end producere.

"The assumption is that at some level or levels of the production process there exist neural signals standing in one-to-one correspondence with the various segments of the language – phoneme, word, phrase, etc. Perception consists in somehow running the process backward, the neural signals corresponding to the various segments being found at their respective levels. In phoneme perception – our primary concern in this paper – the invariant is found far down in the neuromotor system, at the level of commands to the muscles" (Liberman, Cooper, Shankweiler & Studdert-Kennedy 1967: 454).

Denne lineære perceptionsmodel er nok den mest kendte af Motorteoriens varianter og er oftest den model, der refereres til med betegnelsen 'the Motor Theory of Speech Perception'[6]. Motorteorien postulerer i denne form, at perceptionen af tale medieres af viden om artikulationsprocessen. Talen skulle altså ikke perciperes ved analyse af det auditive signal, men derimod ved en speciel proces, som giver lytteren mulighed for at udlede de artikulatoriske kommandoer, der kræves for at producere det akustiske signal.

En vigtig type undersøgelser, der underbygger hypotesen om korrespondens mellem artikulatoriske kommandoer og fonemer, er forsøg med sproglydes oplevede styrke. Lehiste & Peterson konstaterer (1959), at vokaler udtalt med samme anstrengelse er meget forskellige med hensyn til mængden af akustisk energi, f.eks. indeholder [i] meget mere akustisk energi end [o] på trods af, at talere oplever dem som udtalt med samme styrke. De præsenterede vokaler med ens mængde akustisk energi for en række forsøgspersoner og bad dem vurdere deres relative styrke. Konsekvent vurderede forsøgspersonerne, at [o] var meget kraftigere end [i] på trods af en ens mængde akustisk energi i de to stimuli. Med andre ord, konkluderer Lehiste & Peterson, vurderer forsøgspersonerne, at vokalerne har samme styrke, ikke når de består af samme mængde akustisk energi, men når de antages at blive udtalt med samme artikulatoriske anstrengelse. Lytterens model af talerens taleorganer og deres artikulatoriske dynamik er altså helt essentiel; perceptionen knytter sig snarere til antagelser om artikulatoriske forhold hos taleren end det faktiske akustiske signal, som antagelserne udspringer af.

En anden vigtig type materiale er undersøgelser som viser, at perceptionen betinges af lytterens forventninger, baseret på hans viden om artikulations(u)muligheder (Borden & Harris 1984). Med 'forventninger'

[6] Til tider betegnes denne variant også 'the Motor Command Theory' (se f.eks. MacNeilage 1991).

refererer jeg altså udelukkende til forventninger om artikulatoriske forhold og ikke til betydningsorienterede forventninger, der udspringer af top-down-analyser. Det skal pointeres, at denne Motorteori bygger på en lineær model udelukkende med bottom-up-analyser. Et interessant forsøg udført af Dorman, Raphael og Liberman (1979) viser, at lyttere ikke perciperer lydkombinationer, som ikke kan være produceret af menneskelige taleorganer. Ved præsentationen af [ɛbdɛ] kunne forsøgspersonerne ubesværet identificere lukket i [b] og eksplosionen i [d]. Men når de to klusiler blev rykket tættere sammen – så tæt, at det ikke er muligt at producere med menneskelige taleorganer – var forsøgspersonerne ikke længere i stand til at identificere det første lukke, men hørte den syntetiske stimulus som [ɛdɛ]. I en variant af samme forsøg havde Dorman, Raphael og Liberman produceret [ɛb] med én stemme og [dɛ] med en anden stemme. I dette tilfælde identificerede forsøgspersonerne [ɛbdɛ] uanset hvor tæt de to stavelser var rykket sammen. Lytterens forventninger har en stor betydning for perception, ikke mindst de forventninger, der knytter sig til lytterens viden om artikulationens muligheder og umuligheder.

Senere formuleringer af Motorteorien (Liberman & Mattingly 1985) postulerer, at invariante relationer udelukkende eksisterer på det artikulatoriske niveau. Denne ekstreme formulering af Motorteorien insisterer på, at en én-til-én relation udelukkende eksisterer mellem lingvistiske elementer og motoriske kommandoer. Den påstand stemmer dog ikke særlig godt overens med data fra mange års undersøgelser (Lieberman & Blumstein 1988).

Ifølge Motorteorien fortolkes det akustiske signal ved hjælp af motoriske kommandoer, der er nødvendige for at producere et tilsvarende akustisk signal. Da artikulatorisk invarians i produktionen af et bestemt fonem ikke nødvendigvis resulterer i akustisk invarians, kan man med denne teori forsøge at undslippe 'the acoustic-phonetic noninvariance problem'. Men det har endnu ikke været muligt at finde direkte korrespondens med fonemet, hverken i artikulatoriske data eller i akustiske data. Man har altså haft mindst lige så ringe held med at søge fonemrepræsentationer i diskrete muskelaktiviteter som i diskrete akustiske forandringer. En elektromyografisk undersøgelse (MacNeilage 1970) demonstrerer tydeligt, at det ikke er muligt at finde artikulatorisk invarians ved produktionen af fonemer. Man havde

forventet at kunne afdække en én-til-én relation mellem neuronaktiviteten, taleorganernes muskler og de såkaldte artikulatoriske kommandoer:

"A great merit of the Motor Command Theory [:Motorteorien som den fremtræder i Liberman et al. 1967 – TOM] was that it was testable. Electromyographic studies by myself and others (for a summary of these studies, see MacNeilage, 1972) showed that rather than being invariant, the motor commands reaching the muscles for a given phoneme varied with the coarticulatory demands of both adjacent and nonadjacent segments, with stress levels and with speaking rate" (MacNeilage 1991: 61f).

En ensidig fokusering på det artikulatoriske løser altså ikke invariansproblemet, men forskyder det blot til et "motor-phonetic noninvariance problem" (Warren 1983: 288f).

Konklusionen er altså, at Motorteoriens stærke hypotese må forkastes, for det første, fordi dens præmisser slet ikke stemmer overens med forhåndenværende data, og for det andet, fordi den blot komplicerer den allerede indviklede relation mellem lyd og fonem. Motorteorien kan i bedste fald accepteres i form af en svag hypotese der antager, at motoriske kommandoer omsættes til abstrakte størrelser, vi kan anvende i identifikationen af sproglyde.

'The Revised Motor Theory' (herefter 'Den reviderede Motorteori') er Liberman & Mattinglys (Liberman & Mattingly 1985) reaktion på ovenstående og andre problemer ved Motorteoriens stærke hypotese. For at svække kravet om testbare invariante motoriske kommandoer indfører Liberman & Mattingly en dybdestruktur bestående af 'intenderede fonetiske gestus', hvilke de beskriver således:

"...gestures have a virtue that the acoustic cues lack: Instances of a particular gesture always have certain topological properties not shared by any other gesture...gestures do have characteristic invariant properties, as the motor theory requires, though these must be seen, not as peripheral movements, but as the more remote structures that control the movements. These structures correspond to the speaker's intentions" (Liberman & Mattingly 1985: 23).

Ud fra et rent teoretisk perspektiv er Den reviderede Motorteori en klar forbedring af Motorteorien, først og fremmest fordi den nu kan opstille et generelt, helhedsskabende princip, den fonetiske gestus, der forener talens to

aspekter, produktion og perception. Med afsæt i en underliggende dybdestruktur, der afvikles i produktionssystemet og medierer i perceptionsprocessen, antages Den reviderede Motorteori at kunne integrere såvel visuel og taktil som akustisk information.

Motorteoriens svage hypotese er dog blevet anklaget for at være så abstrakt og generel, at den ikke længere kan forudsige noget om selve perceptionsprocessen[7]. Selvom Motorteorien tilbyder en uddybet begrebsliggørelse af perceptionsprocessen, kan den ikke beskrive de grundlæggende rutiner i informationsprocesseringen, som lytteren gennemgår, når det perciperede signal skal konverteres til invariante fonetiske gestus. Kort og godt tilbyder Motorteorien en begrebsliggørelse, men ingen mekanisme[8].

Motorteorien i dens forskellige varianter er gennem tiden blevet udsat for meget hård kritik. De mest centrale punkter i kritikken er, (1) at Motorteoriens forudsigelser ikke stemmer overens med en række vægtige data; (2) at Motorteorien ikke kan tilbyde nogen reel løsning på problemet om akustisk eller motorisk invarians; (3) at Motorteorien ikke kan give en beskrivelse af grundlæggende rutiner i informationsprocesseringen – eller hvordan de tilegnes; og (4) at Motorteorien præsupponerer, at rækkefølgen i sprogtilegnelse er produktion først og perception senere. På trods af al den kritik, danner Motorteorien alligevel grundlag for den måde, mange ser på sprogperception, på forholdet mellem det akustiske signal og segmenter, og på fonembegrebet i eksperimentel fonetik.

Analysis-by-Synthesis

En anden teori, der inddrager den nære relation mellem perceptionen og mekanismer i taleproduktionen, er Analysis-by-Synthesis-teorien, som først og fremmest repræsenteres ved Kenneth N. Stevens og Morris Halle (Halle & Stevens 1959; Halle & Stevens 1972; Stevens 1960; Stevens & Halle 1967). Analysis-by-Synthesis-teorien blev udviklet parallelt med Motorteoriens tidlige faser (dvs. mens den endnu postulerede en stærk motorhypotese), men udgør snarere et supplement end en total omskrivning af 'the Motor Theory of Speech Perception'. Man vil da også finde de to teorier behandlet under samme overskrift i oversigtsværker om sprogperception og relaterede emner

[7] Dette implicerer ubeklagvis, at teorien ikke længere er falsificerbar.
[8] Liberman (1970) har dog argumenteret for, at transformationen fra lyd til fonem er en grammatisk transformation og derfor er en del af sprogbrugerens medfødte lingvistiske kompetence (Cole 1977).

(se f.eks. Lieberman & Blumstein 1988: 147f; Borden & Harris 1984; Harley 1995).

På samme måde som Motorteorien fastholder Analysis-by-Synthesis fonemet som perceptuel enhed uden samtidig at kræve det realiseret i form af invariante akustiske træk. Analysis-by-Synthesis minder meget om Motorteorien, da lytteren også her formodes at referere til et sæt neurologiske kommandoer, der både indgår i perceptionen og produktionen af tale. Det er med andre ord igen de invariante motoriske kommandoer, der skal forbinde det akustiske signal med fonetiske segmenter.

Ifølge Analysis-by-Synthesis identificerer lytteren det akustiske signal ved at sammenligne det med en syntetiseret psykoakustisk skabelon, der konstrueres ved hjælp af en model over taleorganerne og deres artikulatoriske dynamik. Som i andre synteseteorier, bl.a. Motorteorien, finder identifikationsprocessen sted ved en sammenligning mellem det egentlig perciperede signal og den syntetiserede replika. Hvis sammenligningens udfald er positivt, kan lytteren udlede motoriske kommandoer, der anvendes i segmental analyse, direkte fra den syntetiserede psykoakustiske skabelon. Hvis sammenligningens udfald derimod er negativt på grund af tilstrækkelig stor forskel mellem de to auditive mønstre, justerer lytteren den artikulatoriske model, og dermed også den hypotetiske skabelon, for endnu en gang at sammenligne med det perciperede signal (Lieberman & Blumstein 1988).

Betegnelsen for denne perceptionsteori, Analysis-by-Synthesis, eller på dansk 'analyse ved syntese', giver altså en ganske god idé om hovedtrækkene i perceptionsmodellen: Lytteren analyserer det akustiske signal ved sammenligning med en syntetiseret auditiv replika (Halle & Stevens 1959).

Nu er det ikke sådan, at lytteren syntetiserer en psykoakustisk replika ud fra et tilfældighedsprincip totalt uafhængig af det akustiske signal. Der er snarere tale om et 'kvalificeret gæt', som bygger på en (spektral-)analyse af det neurale spektrogram, der udgør det akustiske signals auditive manifestation. Det 'kvalificerede gæt', den hypotetiske replika, gennemgår herefter en trial-and-error-procedure, hvor lytteren forsøger at matche replikaen med det perciperede signal. Det er vigtigt at pointere, at trial-and-error-proceduren er en nødvendighed, hvis modellen skal give mulighed for at minimere forskelle mellem de sammenlignede størrelser. Et af de største problemer ved

Motorteorien er netop, at sprogproduktion er en forudsætning for sprogperception; lytteren skal selv kunne producere en bestemt sproglyd for at kunne identificere den. Analysis-by-Synthesis er derimod en langt mere dynamisk model. Modellen bygger stadig på en bottom-up-analyse, men analysen er blevet mere sofistikeret med en rekursiv trial-and-error-procedure.

Det er en vigtig fordel ved Analysis-by-Synthesis, at modellen tager højde for både inter- og intraindividuel variation. Den rekursive trial-and-error-procedure giver mulighed for, at de psykoakustiske hypoteser kan justeres, så de tilpasses en lang række potentielle talehastigheder, tonehøjder, stemmekvaliteter, udtalenormer, talehandikap, etc. Et ganske kort øjebliks tale er tilstrækkeligt til, at lytteren kan justere hypotesen og den syntetiserede skabelon, så den svarer til talerens særlige taleproduktion. På denne måde kan alle vigtige fonetiske og fonologiske distinktioner på det pågældende sprog fastholdes, samtidig med at analysen tillader netop de særlige kvaliteter, der kendetegner en given taler på et givet tidspunkt og sted.

Ladefoged & Broadbent (1957) har i et forsøg vist i hvor stor udstrækning lytteren tilpasser identifikationsprocessen til talerens artikulatoriske karakteristika. En række forsøgspersoner blev præsenteret for én af seks versioner af en syntetisk fremstillet ytring, 'Please say what this word is', efterfulgt af ét af fire syntetisk fremstillede testord, [bit], [bɪt], [bɛt] og [bæt]. Forsøgspersonernes opgave var så at identificere testordet som ét af de fire mulige. Den korte introduktionsytring, der udgjorde forsøgets varierende parameter, viste sig at have en signifikant indvirkning på identifikationen af testordet. Når den introducerende ytring blev præsenteret med en relativt høj tonehøjde, identificerede forsøgspersonerne den første stimulus som 'bit' i 87 procent af tilfældene. Når den samme introducerende ytring derimod blev præsenteret med en relativt lav tonehøjde, identificerede forsøgspersonerne nu pludselig den samme stimulus som 'bet' i 90 procent af tilfældene. Ladefoged & Broadbents resultater viser, at forsøgspersonerne anvender den introducerende ytring til at bestemme karakteristika ved "talerens" sprogproduktion, og at de justerer deres perception, så den tilpasses disse karakteristika. Resultaterne af dette forsøg er med andre ord helt i overensstemmelse med forudsigelser baseret på Analysis-by-Synthesis.

Den absolut vigtigste pointe ved Analysis-by-Synthesis – set som teoretisk model – er, at den inddrager sprogsystemets mere globale egenskaber. For det første foretages syntesen under hensyn til det pågældende sprogs fonetiske og fonologiske struktur. For det andet syntetiseres segmenter i forhold til ytringens suprasegmentale struktur. Dette er vigtigt, da analyser viser, at relationen mellem det akustiske og det fonetiske niveau varierer med trykforhold i stavelsen, intonationskonturen, samt stavelsens interne og eksterne struktur. Endelig kan det argumenteres, at Analysis-by-Synthesis ved lytterens rekursive analyse er underlagt det enkelte sprogs syntaktiske og semantiske restriktioner (Harley 1995; Borden & Harris 1984; Clark & Clark 1977).

2.1 Introduktion til kategorisk perception

Menneskets evne til at kategorisere er en absolut fundamental kognitiv egenskab. Vi er afhængige af at kunne gruppere genstande, oplevelser, følelser etc. i en ellers kaotisk verden. Når vi kategoriserer, fokuserer vi på vigtige egenskaber, som forskellige objekter har til fælles, og ignorerer alle irrelevante detaljer.

Nogle af de kategorier, der danner grundlaget for vores oplevelse af verden, er naturlige kategorier; de afspejler fremtrædende fysiske forskelle mellem ting i omgivelserne og efterlader ingen tvivl om, hvad der er inkluderet i kategorien og hvad der ikke er[1]. Andre kategorier er langt mindre transparente og afspejler først og fremmest kulturspecifikke konventioner. Hvad enten kategorierne er naturlige eller arbitrære, tænker vi sjældent på dem som kategorier – og slet ikke på det forhold, at de kun eksisterer i vores bevidsthed (Madsen 1996).

Enhver form for kategorisering dækker over to typer adfærd: den ene er diskrimination, det forhold, at vi opfatter A og B som forskellige; den anden er identifikation, det forhold, at vi oplever flere A'er som tilfælde af samme 'A (kategorien "A"). Diskriminationen har altså at gøre med ens/forskellige-vurderinger, hvor identifikationen er en vurdering af, om et element tilhører en bestemt kategori eller ej.

Typisk er vores evne til at diskriminere langt bedre end vores evne til at identificere. Alle, der har prøvet at løbe tør for vægmaling ved, at det er svært at identificere 'havblå' i den lokale malerforretning, men alt for let at se forskellen, når man har malet de sidste to vægge i stuen. Når det vedrører perceptionen af sproglyde, er situationen en helt anden. Her viser gentagne forsøg, at man almindeligvis ikke kan skelne mellem flere sproglyde, end man kan identificere. Den kategoriske perception af sproglyde er enestående, da man i perceptionen af ikke-sproglige lyde almindeligvis kan skelne mellem mange flere stimuli end man kan identificere[2].

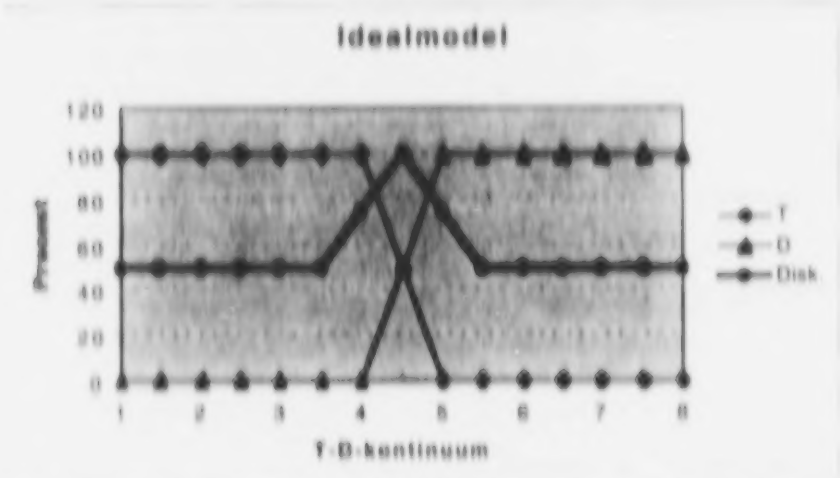
[1] Hvorvidt naturlige kategorier overhovedet har nogen realitet er en gammel filosofisk diskussion, der igen er livet op i løbet af 1980'erne (jf. f.eks. Lakoff 1987; Johnson 1987). Brent Berlin & Paul Kays (1969) opdagelse af, at sanseorganernes biologiske opbygning (receptorernes egenskaber) determinerer oplevelsen af prototypiske farver, viser, at ikke alle kategorier er arbitrære – men om man kan tale om naturlige kategorier, ønsker jeg ikke at diskutere nærmere her, da det ikke har nogen relevans for denne undersøgelse.

[2] Der er dog en række undtagelser, som jeg vil vende tilbage til senere.

Det fænomen, at diskriminationen mellem akustisk ækvidifferente (lige forskellige) stimuli er ringe inden for den sproglige kategori (f.eks. 'de lyde er alle [b]'), men god mellem forskellige sproglige kategorier (f.eks. 'den første var [b], men den næste var [p]') vil herefter blive betegnet 'kategorisk perception' (forkortes 'KP').

'Perception is said to be categorical when subjects are able to discriminate among a set of stimuli only to the extent that they can assign different labels to them' (Studdert-Kennedy et al. 1970).

I nedenstående hypotetiske forsøg med kategorisk perception præsenteres forsøgspersoner for en serie tilfældigt ordnede stimuli fra et kontinuum, f.eks. et [d]-[t]-kontinuum. Der er ofte tale om syntetisk genererede stimuli, der adskiller sig fra hinanden ved lige store skridt ad et akustisk kontinuum. Forsøgspersonerne[1] bedes identificere de præsenterede stimuli, i dette eksempel som 'd' eller 't', hvorved det akustiske kontinuum inddeles i to kategorier, jf. figur 2.1.1.



Figur 2.1.1 Idealiserede identifikations- og diskriminationsfunktioner, der viser kategorisk perception af stimuli fra et [t]-[d]-kontinuum bestående af 8 stimuli.

[1] Nogle steder i teksten vil 'forsøgspersoner' være forkortet 'fp' (især i modeller). Den varierende brug af hhv. den forkortede og den uforkortede form er ikke udtryk for forskellig reference, men udgør alene stilistisk variation.

Dernæst præsenteres forsøgspersonerne for en tilfældigt ordnet serie stimuli i par. De skal nu vurdere, om de pågældende stimuli i hvert par er ens eller forskellige, altså om de kan diskriminere mellem stimuli.

Identifikationsøvelsen afslører typisk en fonetisk grænse – her mellem stimulus 4 og 5. Stavelser til venstre for denne grænse tilskrives konsekvent kategori 'T, mens stavelser til højre for grænsen identificeres som medlemmer af kategori 'D. Modellen forudsiger, at når diskriminationsevnen testes, er den maksimal mellem de to kategorier 'T og 'D, hvorimod diskrimination mellem stimuli fra samme fonetiske kategori er tilfældig, altså omkring 50 procent. Selvom de akustiske forskelle mellem stimulus 1 og 2, 4 og 5, 7 og 8 er de samme, er det ikke muligt at høre, at der er tale om forskellige sproglyde i første og sidste par, men relativt let at høre forskellen mellem stimulus 4 og 5:

"Other things being equal, stimuli belonging to the same phonetic category are more difficult to discriminate than stimuli on opposite sides of a phonetic boundary. This phenomenon has long been known as 'categorical perception'" (Studdert-Kennedy et al. 1970).

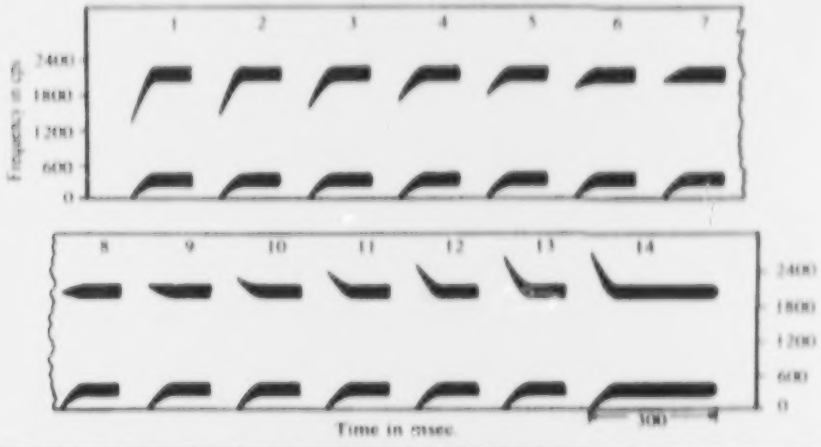
Senere er man dog begyndt at skelne mellem det samlede fænomen 'kategorisk perception'[4] og den specifikke empiriske effekt, som forsøg med kategorisk perception ofte afføder, nemlig en signifikant stigning i diskriminationen ved kategorigrænser (Repp 1982). Denne sidste effekt betegnes almindeligvis 'the phoneme boundary effect' (herefter 'den kategoriske effekt') og blev først beskrevet indgående af Wood i 1975.

Som det fremgår af ovenstående, vedrører forskning i KP de fonetiske segmenters perceptuelle realitet, dvs. fonetiske kategoriers funktion i perceptionsprocessen uafhængigt af, om sprogbrugeren er bevidst om kategoriernes realitet. Selvom KP oprindeligt begyndte som et meget snævert forskningsområde, er det med tiden blevet en betegnelse for et relativt bredt forskningsparadigme med veldefinerede empiriske modeller. I det følgende er det min hensigt at give et historisk og metodisk overblik over netop dette paradigme.

[4] Kategorisk perception dækker over et fænomen, der inkluderer skarpe kategorigrænser, reduceret kontekstsensibilitet, forudsigelig diskriminationsadfærd på baggrund af identifikationsscoren o.a. (Studdert-Kennedy, Liberman, Harris & Cooper 1970). Den teoretiske model vil blive beskrevet nærmere i det følgende.

Kategorisk perception – en historisk og metodisk oversigt

Den tidlige forskning i KP udvikledes ved Haskins-Laboratorierne i New Haven, Connecticut, i takt med brugen af talesynthesizere – også kaldet 'Pattern Playback'[5]. Liberman, Harris, Hoffman og Griffith (1957) skulle blive de første til at beskrive det fænomen, som senere er blevet kendt under betegnelsen 'Kategorisk perception'. Ved hjælp af en Pattern Playback-synthesizer havde de konstrueret et fysisk kontinuum af CV-stavelser, der dækkede de tre kategorier [ba], [da] og [ga]. Selvom disse syntetiske stimuli udgjorde et akustisk kontinuum – fremkaldt ved en gradvis forhøjelse af F2's begyndelsesfrekvens (jf. figur 2.1.2) – grupperede forsøgspersonerne dem i tre skarpt adskilte kategorier.



Figur 2.1.2 Akustisk kontinuum bestående af 14 syntetiske stimuli med varierende F2. I en identifikationstest blev de 14 formantpar kategoriseret som hhv. [ba], [da] og [ga]. (Fra Liberman et al. 1957:359)

For at undersøge om de ekvidifferente fysiske forskelle mellem stimuli fra samme kategori kunne opfanges af forsøgspersonerne, antog Liberman et al. en ABX-diskriminationstest. I denne test skal forsøgspersonerne afgøre, om den sidste af tre konsekutive stimuli er identisk med den første eller anden, der altid er forskellige. Resultaterne af testen viste, at stimuli, der tidligere var blevet identificeret som tilhørende forskellige sproglige kategorier, var

[5] Pattern Playback-talesynthesizeren er resultatet af en forskningsorienteret videreudvikling af den føromtalte Vocoder-synthesizer, der oprindeligt blev udviklet ved Bell Telephone Laboratorierne i slutningen af 1930'erne.

lette at diskriminere, hvorimod stimuli, der tidligere var blevet identificeret som tilhørende samme sproglige kategori, var meget svære at diskriminere. Da det akustiske kontinuum bestod af ækvidifferente stimuli, der var sammenlignelige, kunne resultaterne kun tilskrives perceptionsprocessen.

Med afsæt i en antagelse om, at lyttere ikke har information ud over den fonologiske kategori[6], kunne Liberman et al. (1957) opstille en forudsigelse om forsøgspersoners diskriminationsevne determineret af deres adfærd i identifikationstesten. Imidlertid viste diskriminationstesten, at forsøgspersonerne havde lettere ved at høre forskel mellem (både inter- og intrakategoriale) stimuli end forventet, hvilket kan tolkes derhen, at lyttere benytter sig af yderligere information om stimulus.

Ovenstående forsøg rejste en række centrale spørgsmål om perceptionen af sproglyde og affødte et helt nyt forskningsparadigme. Med Liberman et al. (1957) blev der sat standard for en række lignende undersøgelser med forskellige fonetiske kontraster. F.eks. rapporterede Liberman, Harris, Kinney & Lane (1961) kategorisk perception af [d] versus [t] kontrast ved 'F1 cutback'[7]; Liberman, Harris, Eimas, Lisker & Bastian (1961) fandt lignende resultater med intervokalisk [b] versus [p] kontrast ved varierende 'silent interval'; Bastian, Eimas & Liberman (1961) demonstrerede, at artikulatorisk lukke signaleret ved silent interval (f.eks. [slit] versus [split]) ligeledes blev perciperet kategorisk.

Disse forsøg står i kontrast til forsøg udført af bl.a. Fry, Abramson, Eimas & Liberman (1962) og Eimas (1963), der viser, at syntetiske vokaler i et [I]-[ɛ]-[æ]-kontinuum diskrimineres lige godt i og mellem fonetiske kategorier – et fænomen, der almindeligvis betegnes 'kontinuerlig perception'. Kontinuerlig perception er også blevet fremkaldt med andre akustiske kontinua, f.eks. vokallængde (Bastian & Abramson 1964) og intonationskonturer (Abramson 1961), samt med ikke-sproglige stimuli, der har træk til fælles med

[6] Denne antagelse udspringer af Motorteoriens hypotese om, at det akustiske signal og fonetiske segmenter i både perception og produktion medieres af motoriske kommandoer (jf. side 22).

[7] 'F1-cutback' var den første teknik, der blev anvendt til at skabe et stemthedskontinuum. Teknikken består i en gradvis forsinkelse af F1's onset i forhold til de andre (øvre) formanters onset. 'F1-cutback' er senere blevet afløst af en anden teknik, hvor en forsinkelse af F1's onset kombineres med en substitution af aperiodisk med periodisk energi i de øvre formanter under forsinkelsesperioden. Den sidstnævnte teknik blev udviklet af Abramson & Lisker i løbet af 1970 (Abramson & Lisker 1970; Lisker & Abramson 1970).

kategorisk perciperede sproglige stimuli (f.eks. Liberman, Harris, Eimas, Lisker & Bastian 1961; Liberman, Harris, Kinney & Lane 1961). Med andre ord synes KP udelukkende at vedrøre sproglige stimuli – med undtagelse af isolerede vokallyde – og især klusiler.

Resultaterne fra disse tidlige undersøgelser af KP havde stor indflydelse på udviklingen af perceptionsteorier i løbet af 1950'erne og 1960'erne, især synteseteorierne, hvoraf Motorteorien er den mest kendte (jf. side 20ff)[8]. Det essentielle ved synteseteorierne er, at de bygger på en hypotese om, at sprog-perception og -artikulation bygger på samme (eller nært forbundne) neurologiske processer, dvs. at sprogbrugere trækker på ét sæt neurologiske kommandoer, der både indgår i perceptionen og produktionen af tale. Når forskellige fonetiske kategorier (f.eks. [t]-[d]) adskiller sig fra hinanden ved diskrete artikulatoriske egenskaber (f.eks. artikulationssted (F2) eller artikulationsmåde (VOT[9])), vil et akustisk kontinuum, der strækker sig over flere af disse kategorier, blive perciperet kategorisk. Derimod vil perceptionen være kontinuerlig, når kontinuerlig artikulatorisk variation mellem fonetiske kategorier er mulig (som f.eks. ved [i]-[e]-[æ]-[a]). Både Analysis-by-Synthesis og Motorteorien tager altså KP som en direkte afspejling af den artikulatoriske organisering.

Bortset fra et enkelt kritisk indlæg af Harlan Lane i 1965 (Lane 1965) var forskningen i KP i 1950'erne og 1960'erne udelukkende forbundet med gruppen af forskere ansat ved Haskins-Laboratorierne – og havde udelukkende Motorteorien som teoretiske indgangsvinkel.

Harlan Lane og hans kollegaer ved Michigan Universitet studerede KP ud fra et psykofysisk perspektiv og forsøgte at påvise, at den kategoriske effekt også kan fremkaldes ved hjælp af ikke-sproglige stimuli under sammenlignelige eksperimentelle forhold. Resultatet af en række forsøg blev samlet i et kraftigt udfald mod især forskningen i KP på Haskins-Laboratorierne, men

[8] De fleste forsøg med KP blev udført på Haskins-Laboratorierne, der samtidig var centrum i udviklingen af Motorteorien. Der er ingen tvivl om, at opdagelsen af den kategoriske effekt virkede forstærkende på denne udvikling. Den primære årsag til at der ikke blev udført flere forsøg med KP andre steder end på Haskins-Laboratorierne er simpelthen, at man på dette tidspunkt kun havde adgang til de nye forskningsorienterede talesynthesizere ganske få steder.

[9] VOT (forkortelse af Voice Onset Time) udgør her den tidslige afstand mellem en stavelsesinitiale konsonants eksplosion og stemthedens påbegyndelse. VOT dækker med andre ord eksplosionen og aspirationen. Jeg er klar over, at VOT kan måles fra implosionen, men dette ville ikke give information om eksplosionen og aspirationens varighed (Springer 1980).

også mere generelt de nye synteseteorier (Lane 1965). Der er nok ingen tvivl om, at kritikken i Lanes indlæg foregreb nogle af de alternative forklaringer på KP, der er blevet præsenteret fra midten af 1970'erne og frem til i dag. Lanes indlæg fik dog en relativt ringe indflydelse på samtidens forskning i KP, hvilket først og fremmest skyldtes et svagt empirisk fundament. Faktisk blev Lanes indlæg totalt ignoreret indtil 1970, hvor Haskins-gruppen, repræsenteret ved Studdert-Kennedy, Liberman, Harris og Cooper, tog kraftigt til genmæle (Studdert-Kennedy et al. 1970).

Forskningen i KP fortsatte på Haskins-Laboratorierne op gennem 1960'erne, bl.a. med Abramson & Lisker (1970), der påviste kategorisk perception af stemte versus ustemte lukkelyde (f.eks. [ba]-[pa]) i VOT-kontinua. Forsøgspersoner, der havde thai som modersmål, havde én stemthedsgrænse (eller VOT-grænse) og forsøgspersoner med engelsk som modersmål en anden. En anden tidlig komparativ undersøgelse blev udført af Stevens, Liberman, Öhman & Studdert-Kennedy (1969) med engelske og svenske vokallyde. Selvom resultatet af denne test ikke viste samme grad af kontinuerlig perception som føromtalte undersøgelse af Fry et al. (1962, jf. side 35), var der ingen korrelation mellem identifikation og diskrimination, hvilket udelukker kategorisk perception. Endelig blev de tidlige forsøg med kategorisk perception af f.eks. de stemte lukkelyde [ba]-[da]-[ga] (Liberman et al. 1957) reproduceret i en serie nye undersøgelser, bl.a. en undersøgelse udført af Mattingly, Liberman, Syrdal & Halwes (1971), der som de første inddrog lukkelyde i final position samt en række ikke-sproglige kontrolstimuli, der ikke blev perciperet kategorisk.

Det var dog ikke kun vokal- og lukkelyde, der blev anvendt i forsøg med KP. Efterhånden som talesynthesizerne blev lettere tilgængelige og syntese-teknikkerne mere sofistikerede, blev også andre fonetiske kategorier udforsket. F.eks. blev der afsløret en delvis kategorisk perception af affrikat-frikativ distinktionen (Cutting & Rosner 1974), af likvid-kontinua (McGovern & Strange 1977; Miyawaki, Strange, Verbrugge, Liberman, Jenkins & Fujimura 1975), af nasalkonsonanter (Larkey, Wald & Strange 1978; Miller & Eimas 1977) og af oral-nasal distinktionen (Miller & Eimas 1977).

Informationsprocessering og 'the dual-process model'

Umiddelbart virkede det som om al forskning i KP i 1950'erne og 1960'erne foregik på Haskins-Laboratorierne. Dette viste sig ikke at være tilfældet. I årene 1968-1971 havde to japanske forskere, Fujisaki og Kawashima, beskæftiget sig med KP som informationsprocessering – et arbejde, der var aldeles ukendt for vestlige forskere, indtil Pisoni (1971, 1973, 1975) diskuterede og perspektiverede det.

Med Fujisaki og Kawashima blev fokus nu flyttet fra kategorisk perception som fænomen til kategorisk perception som en perceptionsproces[10]. Fujisaki & Kawashima (1969, 1970, 1971) formulerede, hvad de betegnede som en 'dual-process model' for diskrimination af sproglyde. Modellen skelner eksplicit mellem kategorisk fonemisk perception og auditiv perception, der er knyttet til en (præ-)auditiv buffer, der kan fastholde akustiske træk[11]. Hensigten med modellen er at redegøre for de observerbare afvigelser fra den idealiserede Haskins-model – at diskriminationsadfærden i konkrete forsøg er bedre end forventet.

Fujisaki og Kawashima anvendte endvidere nye serier af sproglyde, bl.a. syntetiske frikativer, semivokaler og likvider, og viste, at disse perciperes mindre kategorisk end klusiler, men heller ikke kontinuerligt som isolerede vokallyde. Endelig eksperimenterede Fujisaki og Kawashima med vokallyde af varierende længde og med varierende kontekst, hvilket viste, at selv vokaler kan perciperes relativt kategorisk, hvis forsøgsopstillingen er ugunstig for den auditive hukommelse.

Pisoni (1971, 1972, 1973, 1975, Pisoni & Lazarus 1974) afprøvede og udvidede Fujisaki og Kawashimas 'dual-process model' ved en række forsøg, der bl.a. afslørede, at graden af kontinuitet i perceptionen af sproglyde i høj grad afhænger af, om lytteren er i stand til at anvende den auditive hukommelse. Pisoni cementerede yderligere denne opdagelse i forsøg med varierende stimuluslængde, varierende pause mellem stimuluspræsentationer og ved introduktion af distraktionslyde mellem stimuli, der skulle

[10] Dette medførte bl.a. en større interesse for forskellige forsøgsvariabler, f.eks. stimulusvariabler.

[11] Den auditive buffer/hukommelse – også kendt som 'Preperceptual Auditory Storage' – er en mekanisme, der er blevet beskrevet indgående af psykologer og lingvister, bl.a. Massaro (1970, 1971, 1972), Massaro & Oden (1980), Oden & Massaro (1978), Pisoni (1971, 1972, 1973, 1975), Woosl (1975), etc. På det neurale niveau består den auditive buffer af ganglieceller i nervus acusticus (i det cortiske organ).

diskrimineres. Et forsøg med reaktionstid som indikator for forsøgspersoners (interfonemiske) diskriminationsevne (Pisoni & Tash 1974) satte ligefrem gang i en stigende interesse for forsøgspersoners evne til subfonemisk diskrimination af akustiske forskelle[12]. Således ændrede forskningen i KP i løbet af få år karakter fra udelukkende at skabe fundament for en synteseteori til også at rejse en række spørgsmål inden for en psykoakustisk og psykofysisk beskrivelsesramme.

Kategorisk perception i et nyt perspektiv

I løbet af 1970'erne udspaltedes en række nye semi-uafhængige forskningsparadigmer fra den traditionelle forskning i KP. Et af de nye projekter var udforskningen af spædbørns sprogperception (for en oversigt se Eimas 1975b og Morse 1978). Det eksperimentelle paradigme adskiller sig ikke væsentligt fra forsøg med voksne forsøgspersoner, bortset fra responsprocedurerne High-Amplitude Sucking og Heart Rate[13].

I en af de første og mest berømte undersøgelser af spædbørns sprogperception anvendte Eimas, Siqueland, Jusczyk & Vigorito (1971) HAS-proceduren[14] for at beskrive 1 og 4 måneder gamle spædbørns perception af CV-stavelser ([ba]-[pa]) fra et VOT-kontinuum.

I en test med intrakategorial stimulus-skifte blev spædbørnene først præsenteret for én stavelse – enten med +20 ms eller +40 ms VOT – hvorefter der skiftedes til den anden efter en tilvænningsperiode[15]. En tilsvarende procedure blev anvendt ved interkategorial stimulus-skifte, hvor spædbørnene blev præsenteret for forskellige stimuli fra samme kategori: -20

[12] Mens forskningen i KP udelukkende fandt sted på Haskins-Laboratorierne, der var centrum i udviklingen af Motorteorien, blev enhver tale om subfonemisk diskrimination – af gode grunde – helt undgået, medmindre diskriminationsprocenten var stor nok til at det kunne karakteriseres som 'kontinuerlig perception' (af f.eks. vokallyde) (Repp 1984).

[13] High-Amplitude Sucking og Heart Rate forkortes hhv. HAS og HR. Se Morse 1974 for en diskussion om de to teknikker.

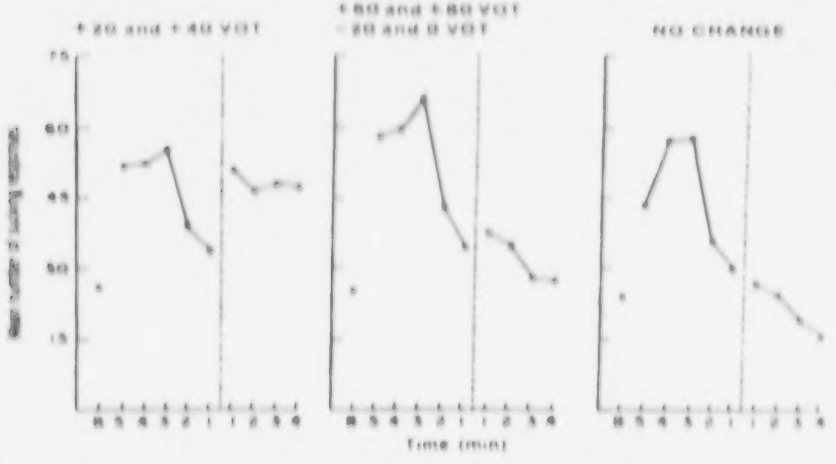
[14] Med HAS-proceduren er præsentationen af en sproglig stimulus betinget af, hvor hurtigt spædbarnet sutter på en elektronisk trykfølsom narresut: Ved præsentationen af en ny stimulus stiger suttehastigheden fra et ellers stabilt niveau, indtil lyden mister sin "nyhedsværdi" og suttehastigheden atter falder til det stabile niveau. En stigning i suttehastighed umiddelbart efter præsentationen af en ny stimulus tolkes i dette paradigme som konsekvens af barnets evne til at skelne mellem to stimuli, den nye og den gamle. Eimas (1975b) har dog pointeret, at en præsentationsprocedure som HAS (og HR) bestemt ikke er uproblematisk: der kan være god grund til at antage, at selve proceduren fremkalder en stimulusadaption (se diskussion af 'selective adaptation' side 44 her).

[15] Denne præsentation er intrafonemisk, da den fonetiske [b]-[p]-kategorigrænse på engelsk er omkring +25 ms VOT (Springer 1979).

BEST COPY AVAILABLE

ms VOT og 0 ms VOT, der begge var blevet identificeret som [ba] af voksne forsøgspersoner, eller +60 ms VOT og +80 ms VOT, der begge var blevet identificeret som [pa], ligeledes af voksne forsøgspersoner. En kontrolgruppe spædbørn blev udelukkende præsenteret for én af de 6 CV-stavelser og havde ingen signifikant uregelmæssig suttehastighed.

Resultaterne viste, at spædbørn, der var blevet præsenteret for interkategorial stimulus-skifte, havde reageret ved en lille stigning i suttehastighed, hvorimod spædbørn, der var blevet præsenteret for intrakategorial stimulus-skifte, udviste en adfærd parallel med kontrolgruppen. Eimas et al. (1971) tolkede disse resultater som et udslag af spædbørns evne til kategorisk perception svarende til voksnes: spædbørn er i stand til at diskriminere mellem forskellige stimuli fra forskellige sproglige kategorier, men ikke akustisk forskellige stimuli fra samme kategori (i det mindste ikke i sådan en grad, at det påvirker suttehastigheden). Netop dette forsøg er blevet udvidet til også at dække andre sproglige kontinua, almindeligvis med resultater svarende til ovenstående (Eimas 1974a, 1974b, 1975a).



Figur 2.1.3 Gennemsnitlig suttehastighed hos 4 måneder gamle spædbørn i 3 forskellige forsøgsopstillinger. 'B' repræsenterer det stabile niveau før præsentationen af en stimulus. Hvert diagram viser suttehastigheden som funktion af tid. Den stiplede linje registrerer stimulusskifte bortset fra i diagrammet længst til højre, hvor den markerer det tidspunkt, hvor der ville have været stimulusskifte. Stimuli i venstre diagram dækker over et [b]-[p]-kontinuum, som det opleves af voksne forsøgspersoner, hvorimod kontrasterede stimuli i det midterste diagram er fra samme kategori. (Fra Borden & Harris 1984:202 – tilpasset fra Eimas, Siqueland, Juszyck & Vigorito 1971:304).

BEST COPY AVAILABLE

Resultaterne fra forsøg med KP hos spædbørn er ofte blevet anvendt som argument for en medfødt sprogkompetence. Om KP hos spædbørn skyldes medfødt kompetence eller tidlig sprogtilegnelse, skal efter min vurdering undersøges ved komparative studier af den lingvistiske erfarings betydning for spædbørns perception af sproglyde (Lasky, Syrdal-Lasky & Klein 1975; Streeter 1976). De forsøg, der er foretaget inden for dette område, er dog temmelig tvetydige og rejser snarere flere (metodiske) spørgsmål, end de besvarer. Et eksempel er Streeter (1976), som arbejdede i Kenya med spædbørn, hvis forældre talte kikuyu – et bantusprog med kun én bilabial lukkelyd, et præstemt [b] med (i gennemsnit) -64 ms VOT. Efter at have gennemført en diskriminationstest med et VOT-kontinuum og samme HAS-procedure som i Eimas et al. (1971), hævdede hun, at kikuyu-spædbørn ikke alene kan diskriminere den relevante kontrast for deres modersmål, men også den relevante kontrast for engelsk – en kontrast, der vel at mærke ikke optræder i kikuyu. Dette resultat stemmer overraskende (?) godt overens med Eimas (1975b), der antyder, at amerikanske spædbørn kan diskriminere stimulus-skifte over præstemt-stemt-kategorigrænsen, som den manifesteres på kikuyu, men ikke kategorisk som med den engelske stemte-ustemte [b]-[p]-kontrast. Morse (1978) har dog konstateret – efter en nøjere analyse af Streeters resultater – at der simpelthen ikke er evidens i hendes data for en sådan påstand om kikuyu-spædbørns kategoriske perception af den engelske [b]-[p]-kontrast (Springer 1979).

Et andet nyt forskningsprojekt, der udvikledes i løbet af 1970'erne, var udforskningen af KP hos forskellige dyrearter. Dette projekt var en naturlig konsekvens af synteseteoriernes hypotese om en (artsspecifik) sprogspecifik perceptionsmodus: Antagelsen om, at sprogproduktion og sprogperception bygger på de samme neurologiske kommandoer, forudsætter, at det ikke er muligt at finde KP hos dyr, da de ikke evner at producere tale (Liberman 1974).

På trods af et relativt lille antal er forsøg med KP hos dyr virkelig kommet i fokus. Nogle af de mest kendte er Kuhl & Miller (1975, 1978); Kuhl (1976); Morse & Snowdon (1975); Sinnott, Beecher, Moody & Stebbins (1976); Waters & Wilson 1976.

Ved hjælp af en HR-procedure[16] til registrering af diskriminationen fandt Morse & Snowdon (1975) både intra- og interkategorial diskrimination hos rhesusaber, dog med en ringere diskrimination af forskellige stimuli fra samme kategori. Desuden kunne Miller & Morse (1976) rapportere en bedre intrakategorial diskrimination hos rhesusaber end hos 3 måneder gamle spædbørn. Også i denne komparative undersøgelse var den interkategoriale diskrimination betydelig bedre end den intrakategoriale.

Kuhl og Millers (Kuhl & Miller 1975; Kuhl 1976) berømte forsøg med chinchillaer (en lille sydamerikansk gnaver), som blev trænet i kategorisering af et VOT-kontinuum[17], viser en VOT-kategorigrænse ved 33,5 ms VOT – meget tæt på den (engelske) fonetiske kategorigrænse ved 35,2 ms VOT, som forsøg med voksne forsøgs-personer afslører. Kuhl & Miller konkluderer på linje med mange andre forskere inden for dette område: Da det ikke er sandsynligt, at disse resultater skyldes en sprogspecifik fonetisk processering hos chinchillaer, må årsagen findes i chinchillaens (og menneskets) auditive psykofysiske indretning (se Springer 1979 for diskussion af denne forklaringsmodel).

Det er efter min mening tvivlsomt, om denne konklusion er rimelig. For det første blev Kuhl og Millers chinchillaer trænet i CV-identifikation, hvilket leder tanken hen mod den klassiske betingningsmetode, kendt fra bl.a. den tidlige adfærdspsykologi omkring århundredeskiftet. For det andet afslører en isoleret identifikationstest ikke, om der er tale om KP: Selv i kontinuerlig perception af et vokal-kontinuum er vi i stand til at identificere de enkelte vokallyde. Derimod er den kategoriske effekt i en diskriminationsøvelse en ganske god indikator for, om vi opfatter de præsenterede stimuli som diskrete størrelser.

Generelt mener jeg, at der er god grund til at forholde sig kritisk til forsøg med KP, der alene baseres på 'evoked potentials' (f.eks. ved hjælp af en HAS-

[16] HR-proceduren minder meget om den føromtalte HAS-procedure (jf. note 14). Med HR-proceduren er præsentationen af en sproglig stimulus betinget af respondentens hjerteslag, hvadenten respondenten er et dyr, et spædbarn eller et voksent menneske: Ved præsentationen af en ny stimulus stiger antallet af hjerteslag, hvorefter det falder til normalniveauet i løbet af en tilvænningsperiode. En stigning i antallet af hjerteslag lige efter præsentationen af en ny stimulus tolkes her som en konsekvens af respondentens evne til at skelne mellem to stimuli.

[17] Der er tale om et akustisk kontinuum svarende til det, der blev anvendt i Eimas, Siqueland, Jusczyk & Vigoritos (1971) forsøg med spædbørn, bortset fra at der her blev anvendt stemte-ustemte bilabiale lukkelyde og ikke stemte-ustemte alveolære lukkelyde.

eller HR-procedure), dvs. forsøgsparadigmer, der ofte anvendes i forsøg med spædbørn og dyr. Resultater fremkaldt ved subliminal perception er simpelthen ikke relevante i beskrivelsen af sprogperception; Om en forsøgspersons hjerteslag ændres ved præsentationen af en ny stimulus er absolut irrelevant i et forsøg med KP, hvis det ikke på anden vis kan dokumenteres, at forsøgspersonen hører forskel. Når jeg alligevel har valgt at inddrage og beskrive en række forsøg, hvis metodologi og analysekoncept er kritisabelt, skyldes det, at deres konklusioner har haft en forholdsvis stor indflydelse på udviklingen af teorier inden for KP, perceptionspsykologi, neurolingvistik, psykolingvistik og andre relaterede områder.

Kategorisk perception i et psykoakustisk perspektiv

Den tidlige forskning i KP, centreret omkring Haskins-gruppen, var nået frem til den konklusion, at KP er specifik for sproglyde – først og fremmest lukkelydene [p], [t], [k], [b], [d] og [g]. Lanes (1965) forsøg på at forklare KP ud fra generelle auditive processer havde ikke haft den nødvendige gennemslagskraft og fremstod derfor næsten upåagtet indtil midten af 1970'erne, hvor den psykoakustiske forklaringsmodel blev "genopdaget".

Haskins-gruppens tidlige konklusioner vedrørende KP viste sig ikke at være holdbare, da forskere begyndte at melde om KP af ikke-sproglige stimuli. Cutting & Rosner (1974) hævdede, at de havde registreret kategorisk perception af komplekse, periodiske toner fra et såkaldte 'pluck-bow' kontinuum[15]. Miller, Wier, Pastore, Kelly & Dooling (1976) rapporterede om kategorisk perception af summelyde, der skulle ækvivalere et VOT-kontinuum. Pisoni (1977) fik et lignende resultat med periodiske toner, der varierede mht. relativ onset time.

Som en konsekvens af ovenstående undersøgelser begyndte flere og flere at opbygge teoretiske psykoakustiske modeller, der kunne integrere resultaterne fra forsøg med KP: Macmillan, Kaplan & Creelman (1977) forsøgte at beskrive KP ud fra en signal-detektor-teori; Pastore (1981) studerede de psykoakustiske faktorer, der kan have indvirkning på KP; og Schouten (1980, 1987), der antog en auditiv 'threshold model', gik så langt som til at hævde, at

[15] Ved hjælp af en Moog-synthesizer genererede Cutting og Rosner et kontinuum af savtakkede lydsvingninger af en varighed på 1000 ms. Stimuli varierede mht. stigningstidspunkt fra 0 til 80 ms i konsekutive skridt á 10 ms. I en identifikationstest blev de første 3 stimuli identificeret som 'knipset' og de sidste 5 som 'strøget' (musikterminologi lånt fra strygerinstrumenter som f.eks. violin).

sprogperception i det hele taget kunne forklares ved psykoakustiske principper. Fællesnævneren for disse psykoakustiske tilgange er en antagelse om, at lingvistiske kategorier grundlæggende er psykoakustisk determineret (Miller, Wier, Pastore, Kelly & Dooling 1976). Dette synspunkt er i løbet af de sidste dekader blevet et reelt alternativ til Fujisaki og Kawashimas 'dual-process model' (beskrevet side 38f). Endvidere er kontroverserne mellem de to tilgange hyppigere og hyppigere blevet forbundet med den gamle debat om, hvorvidt det er nødvendigt at antage eksistensen af en speciel sprog-processor, hvilket hævdes af synteseteoretikerne.

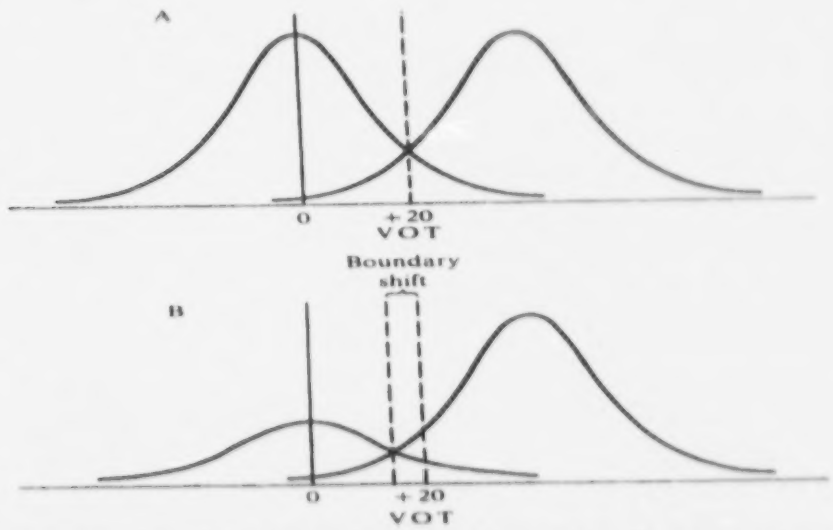
En måde at overflødiggøre den fonetiske sprogprocessor er at foreslå en empirisk model, der kan redegøre for KP ved hjælp af andre teknikker end en decideret fonetisk analyse. Hvis denne model endvidere kan redegøre for andre fænomener, som f.eks. 'dual-process modellen' ikke kan tilgodese, må den anses for at være teoretisk mere robust. Netop denne strategi dannede fundamentet for et forskningsparadigme, 'selective adaptation' der ligeledes blev udspaltet fra KP i løbet af 1970'erne – men senere skulle vise sig at være en metodisk blindgyde.

Man var i løbet af 1970'erne begyndt at overveje den mulighed, at detektorer til identifikation af fonetiske træk var en fundamental mekanisme i perception af sprog (Ades 1976; Studdert-Kennedy 1974)[19]. Eimas & Corbit (1973) var de første til at teste en hypotese, der udsprang af en 'træk-detektor-model'. De postulerede, at VOT-sensitive træk-detektorer var årsagen til stemt-ustemt kontrasten (f.eks. [b]-[p]): en gruppe detektorer reagerer på lav VOT og en anden på høj VOT. Med andre ord ville en stimulus blive identificeret som et stemt fonem, hvis stimulationen af detektorer, der reagerer på lave VOT-værdier, oversteg stimulationen af VOT-detektorer med en høj reaktions-tærskel – og i modsat fald ville stimulus blive identificeret som et ustemt fonem. Stemt-ustemt kategorigrænsen korresponderer ifølge modellen med den VOT-værdi, der virker som lige kraftigt irritament for begge grupper af VOT-detektorer (jf. figur 2.1.4 - A).

Eimas & Corbit (1973) anvendte en teknik, 'selective adaptation', der går ud på at overstimulere (dvs. trætte) den ene af de to grupper VOT-detektorer

[19] Den grundlæggende tanke, at fonemer kan beskrives ved en trækanalyse (feature analysis), var bestemt ikke ny (jf. f.eks. Jakobson, Fant & Halle 1963; Halle 1964; Chomsky & Halle 1968) – det var derimod forsøget på at implementere trækanalysen i en empirisk perceptionsmodel.

ved gentagne gange at præsentere stimuli fra enten den ene eller den anden ende af et VOT-kontinuum. Herefter skal forsøgspersonerne identificere stimuli fra den resterende del af det akustiske kontinuum[20]. I en sammenligning af identifikationsfunktionerne hhv. før og efter 'selective adaptation' fremstår det tydeligt, at den fonetiske kategorigrænse er blevet manipuleret: efter en gentagen præsentation af f.eks. en -10 ms VOT-stimulus er kategorigrænsen blevet rykket mod den stemte ende af det akustiske kontinuum, dvs. at forsøgspersonerne har kategoriseret færre stimuli som stemte end i den tilsvarende præsentation uden 'selective adaptation' (jf. figur 2.1.4 - B). Og akkurat det samme gør sig gældende i diskriminationstesten: diskriminationstoppen er rykket væk fra den ende af det akustiske kontinuum, der korresponderer med de overstimulerede VOT-detektorer. Resultatet af forsøget med 'selective adaptation' svarer fuldstændig til det forventede ifølge 'træk-detektor-modellen'.



Figur 2.1.4 Hypotetiske sensitivitetsfunktioner for VOT-detektorer, der varetager distinktionen mellem f.eks. [b]-[p]. **A**: Før adaption. **B**: Efter adaption med gentagne præsentationer af stimulus. (Fra Foss & Hakes 1978: 91, fig. 3).

Eimas & Corbits (1973) forsøg med 'selective adaptation' er sidenhen blevet gentaget med andre VOT-kontinua (Eimas, Cooper & Corbits 1973; Cooper 1974c). Endvidere er forsøgsrepertoiret blevet udvidet til også at indbefatte

[20] Den gentagene præsentation af stimuli fra en ende af et VOT-kontinuum skete ved en hyppighed af 2 stimuli per sekund i blokke à ét minuts varighed.

forsøg med stimuli, der varierer mht. artikulationssted (Cooper 1974a, 1974b; Ades 1974).

Umiddelbart skulle man nu tro, at sagen var afklaret: 'træk-detektor-modellen kan både forklare KP samt forudsige resultatet af forsøg med 'selective adaptation'. Imidlertid har det senere vist sig, at adaptionseffekten ikke finder sted på et fonetisk niveau, der kan forbindes med 'træk-detektorer', men er et rent auditivt fænomen, der også kan fremkaldes med ikke-sproglige stimuli (Ades 1974, eksperiment 1; Roberts & Summerfield 1981; Sawusch & Jusczyk 1981)[21].

Resultatet af forsøg med 'selective adaptation' er et velkendt fænomen i forsøg med nerveceller – uanset hvor de findes i kroppen og uanset om kroppen tilhører et menneske eller et dyr (Whitfield & Evans 1965). Det neurofysiologiske fænomen betegnes almindeligvis 'the neuron activation threshold'.

Den neurologiske tærskel, der forårsager resultatet i forsøg med 'selective adaptation', er en af de vigtigste funktioner i tilegnelsesprocessen. Når flere nerveceller stimuleres simultant vha. dopamin, der er en neurokemisk transmitter (et opiat), forstærkes synapserne, dvs. forbindelsen mellem neuronerne, og de danner et engram (Hebbs lov). Hukommelsen udgør netop en akkumulation af engrammer. En følge af den forstærkede forbindelse mellem neuronerne er, at neuronernes aktiveringstærskel sænkes, så de kræver en mindre stimulation for at blive aktiveret. Dette er den neurofysiologiske årsag til, at det er muligt at manipulere med en fonetisk kategorigrænse i forsøg med 'selective adaptation'. Som det tydeligt fremgår, har effekten ikke noget med det fonetiske niveau at gøre, men forårsages på et neurologisk og auditivt niveau.

'Træk-detektor-modellen' kan altså hverken forklare resultater af forsøg med 'selective adaptation', KP eller overflødiggøre den fonetiske sprogprocessor. Det grundlæggende problem ved modellen er, at den udgør en tautologi og ikke en forklaring:

"...the error lies in offering to explain phonetic capacity by making a substantive physiological mechanism out of a descriptive property of language" (Studdert-Kennedy 1982: 225). For

[21] For en diskussion om forskellen mellem fonetisk og auditiv adaption, se Springer (1979).

"...the perceived feature is an attribute, not a constituent, of the percept, and we are absolved from positing specialized mechanisms for its extraction" (Studdert-Kennedy 1982: 227).

I dette kapitel har jeg forsøgt at give en beskrivelse af forskningsparadigmet KP. Det har været min hensigt at give et kronologisk overblik over et paradigme, der gennem de sidste 40 år har været forskningsemne for en lang række forskellige fagområder. Denne fagpluralisme, der kendetegner mange grænseområder mellem de etablerede videnskaber, har haft den betydning for forskningen i KP, at udviklingen af metodologien nøje korresponderer med hvem, der har beskæftiget sig med KP og hvornår. Netop for at undgå de mange gentagelser, der ville være konsekvensen af separate afsnit om kronologien og metodologien, har jeg valgt at behandle disse to aspekter i en samlet beskrivelse – naturligvis med fare for at komme til at negligere de "små" enestående, men oplysende undersøgelser. Det har altså været min strategi kun at beskrive, hvad jeg opfatter som de væsentligste bidrag og tendenser, snarere end at tilbyde en kvantitativ gennemgang.

I det følgende er det min hensigt at diskutere en række teoretiske og metodiske spørgsmål, samt beskrive nogle af de vigtigste faktorer, der indgår i forsøg med KP.

2.2 Forsøg med kategorisk perception

Som det fremgår af foregående kronologiske og metodiske gennemgang af forskningsparadigmet kategorisk perception, har mange års undersøgelser af dette fænomen rejst en lang række spørgsmål, men kun besvaret ganske få. Et af de væsentligste spørgsmål er – efter min opfattelse – *hvorfor* visse stimuli i højere grad end andre kan fremkalde en kategorisk effekt, mens andre igen slet ikke fremkalder en kategorisk effekt, men snarere tyder på kontinuerlig perception: resultaterne af en lang række undersøgelser viser, at lukkelydene [p], [t], [k], [b], [d] og [g] fremkalder den kraftigste kategoriske effekt; nasaler, lateraler og frikativer en mindre kategorisk effekt; og at vokaler perciperes kontinuerligt.

Ovenstående spørgsmål om årsagen til asymmetri i KP af forskellige sproglyde kan naturligvis ikke besvares, medmindre man er i stand til at give en beskrivelse af de processer, der er involveret i sprogperception, og som enten forårsager den kategoriske effekt (fortolkes her som resultatet af KP) eller ikke resulterer i en kategorisk effekt (fortolkes her som resultatet af kontinuerlig perception). De eksisterende konkurrerende perceptionsmodeller bygger på antagelser, der spænder fra rent fonetisk sprogperception til rent auditiv sprogperception: Den traditionelle Haskins-model tager udgangspunkt i en hypotese om, at sprogperceptionen er kendetegnet ved fonetisk syntese og at den er neurologisk nært forbundet med sprogartikulation (jf. side 22ff). Den største teoretiske modsætning til Haskins-modellen er nok Schoutens (1980, 1987) 'threshold-model', der postulerer, at (også) sprogperception kan forklares ud fra generelle (dvs. ikke specifikt sproglige) psykoakustiske principper. De to yderpunkter medieres af en række modeller, der antager, at sprogperceptionen delvis trækker på fonetiske processer og delvis på auditive processer – dog med stor forskel på vægtningen af de to perceptionsmodi. Samtidig varierer disse modeller mht. om sprogperceptionen er en aktiv proces (f.eks. dual-process modellen) eller en passiv proces (f.eks. rækken af auditive eller fonetiske træk-detektor-modeller).

Problemstilling

I det følgende vil jeg diskutere 3 hypoteser, der har stor betydning for tolkningen af fænomenet 'Kategorisk perception'. De tre hypoteser er hypotesen om stimulus-asymmetri, motorhypotesen og hypotesen om sprog-specifikke kategorigrænser. Endvidere vil jeg overveje, hvilke betingelser

stimuli i forsøg med KP skal opfylde for at kunne belyse de mest centrale spørgsmål.

Hypotesen om stimulus-asymmetri

Forsøg med KP viser almindeligvis, at lukkelyde (i CV-stavelser) i højere grad end nasaler, lateraler og frikativer fremkalder en kategorisk effekt, mens isolerede vokallyde slet ikke fremkalder en kategorisk effekt. Disse resultater har ført til en generel dikotomisk beskrivelse med perceptionen af hhv. vokaler og konsonanter som to adskilte fænomener.

Data fra empiriske undersøgelser viser, at situationen er langt mere kompleks: For det første viser forsøg med KP sjældent ideel kategorisk perception, dvs. en diskriminationsadfærd, der udelukkende er betinget af kategoriseringen. Oftest er diskriminationen af stimuli fra samme kategori simpelthen bedre end forventet[1]. For det andet viser undersøgelser med vokallyde præsenteret under varierede forsøgsbetingelser[2], at graden af kategorisk effekt kan manipuleres: hvis de præsenterede vokalstimuli er tilstrækkelig korte, indgår i VC-stavelser eller udgør diftonger, kan en kategorisk effekt fremkaldes (dog ikke en kategorisk effekt, der er lige så kraftig som ved præsentationen af lukkelyde).

Der er altså meget, der tyder på, at en dikotomisk beskrivelse af vokal- og konsonantperception ikke er adækvat i forbindelse med forsøg med KP. Det er nok snarere nødvendigt at tale om en graduering af KP, selvom dette strider mod Haskins-gruppens definition:

"Categorical perception refers to a mode by which stimuli are responded to, and *can only be responded to, in absolute terms*. Successive stimuli drawn from a physical continuum are not perceived as forming a continuum, but as members of discrete categories. They are identified absolutely, that is, independently of the context in which they occur. Subjects asked to

[1] Bortset fra Fry et al. (1962), der fandt en noget nær ideel kontinuerlig effekt i diskriminationen af vokallyde, viser alle forsøg, jeg er bekendt med, en bedre diskrimination end forventet. Cross & Lane (1964, citeret hos Lane 1965) har senere anvendt de samme magnetbånd, som Fry et al. (1962) anvendte, men uden samme ideelle kontinuerlige resultat (det skal dog tilføjes, at Cross & Lane tilføjede yderligere resonans, hvilket burde være irrelevant mht. diskriminationen af vokallyde).

[2] Det drejer sig her først og fremmest om 'masking' (bl.a. Massaro 1970, 1971; Repp et al. 1979), interstimulus interval (bl.a. Cutting, Rosner & Foard 1976; Pisoni 1971, 1973; Repp et al. 1979), kontekst (bl.a. Fujisaki & Kawashima 1969, 1970; Pisoni 1975; Repp et al. 1979), rangordning og reaktionstid (bl.a. Oden & Massaro 1978; Perey & Pisoni 1978; Pisoni & Tash 1974; Repp 1975).

discriminate between pairs of such "categorical" stimuli are able to discriminate between stimuli drawn from different categories, but not between stimuli drawn from the same category. In other words, discrimination is limited by identification: subjects can only discriminate between stimuli that they identify differently" (Studdert-Kennedy, Liberman, Harris & Cooper 1970: 234, deres kursiv).

I det følgende forsøg er det min hensigt at efterprøve hypotesen om en signifikant forskel mellem perceptionen af vokal- og konsonantlyde, her præsenteret under forhold, der almindeligvis postuleres at fremkalde den største effekt. Det er min forventning at finde følgende 3 karakteristika i et forsøg med et VOT-kontinuum:

1) Et markant skifte i identifikationsfunktionen.
Dvs. en relativ kraftig stigning eller fald i identifikationsfunktionen. Det punkt på det akustiske kontinuum, der inkluderes lige godt i to kategorier, udgør en kategorigrænse.

2) En markant god diskrimination af stimuli fra forskellige kategorier.
Dvs. en maksimal diskrimination ved kategorigrænser (også kendt som den kategoriske effekt).

3) En markant ringe diskrimination af stimuli fra samme kategori.
Dvs. at stimuli fra samme kategori diskrimineres i færre tilfælde end stimuli fra forskellige kategorier.

I et forsøg med et vokalkontinuum forventer jeg ikke at finde disse træk, der er karakteristiske for KP af lukkelyde. Derimod forventer jeg at kunne registrere følgende karakteristika:

1) Et markant skifte i identifikationsfunktionen.

2) En ringe diskrimination af stimuli, uafhængigt af om de tilhører samme eller forskellige kategorier.

Motor-hypotesen

Ifølge synteseteorierne (jf. Haskins-modellen) udspringer sprogproduktion og sprogperception af samme sæt artikulatoriske kommandoer. Denne antagelse medfører, at sprogbrugere ikke har adgang til information ud over de

fonetiske kategorier: fonetiske kategorier, der adskiller sig ved diskrete artikulatoriske egenskaber, perciperes som absolutte, diskrete størrelser, hvorimod perceptionen er kontinuerlig, når kontinuerlig artikulatorisk variation mellem fonetiske kategorier er mulig. Med afsæt i denne antagelse tager synteseteorierne KP som en direkte afspejling af menneskets artikulatoriske potentiale:

"When different phonetic categories are distinguished by essentially discrete articulatory gestures (as with stop consonants differing in voicing or place of articulation), perception of stimuli from a physical continuum spanning these categories will be categorical; on the other hand, when continuous articulatory variations between phonetic categories are possible (as with the vowels), perception will be continuous" (Repp 1984: 246).

Hvis der eksisterer denne nære sammenhæng mellem artikulation og perception, må det også betyde, at jo mindre artikulatorisk kontinuerlige sproglydene er, desto større kategorisk effekt vil de udvise – og større kontinuitet vil resultere i mindre kategorisk effekt. Denne hypotese er i fuld overensstemmelse med Haskins-modellen og den underbygges af en lang række forsøg med KP: Cutting 1974; Eimas 1963; Fry et al. 1962; Liberman et al. 1961, 1972; Pisoni 1971, 1973, 1975; Stevens et al. 1969, etc.

Det er min hensigt at efterprøve ovenstående hypotese i et forsøg med KP, hvor jeg præsenterer en serie stimuli, der udgør et akustisk kontinuum, men som samtidig er mere eller mindre artikulatorisk kontinuerlig. Hvis artikulatorisk kontinuitet er den eneste faktor, der betinger KP, vil diskriminationen af akustisk diskontinuerlige stimuli være bedre end af akustisk kontinuerlige stimuli. Hvis dette ikke viser sig at være tilfældet, må konklusionen nødvendigvis være, at perceptionen af sproglyde ikke udelukkende bygger på artikulatoriske kommandoer.

Hypotesen om sprogspecifikke kategorigrænser

Et af de mest centrale spørgsmål i forskningen i KP vedrører årsagen til den kategoriske effekt, dvs. årsagen til, at forsøgspersoner udviser en relativ god diskrimination af stimuli fra forskellige kategorier, men ringe diskrimination af stimuli fra samme kategori. I hidtidige forsøg på at give en adækvat beskrivelse har man enten postuleret, at den kategoriske effekt forårsages af en fonetisk eller af en auditiv processering (for en oversigt se Strange & Jenkins 1978).

En af de metoder, man kan anvende i bestræbelsen på at afdække forholdet mellem auditive og fonetiske processer, er komparative forsøg med forsøgspersoner, der taler forskellige modersmål. Med denne metode forsøger man at afsløre, hvordan forsøgspersoner med forskellige modersmål kategoriserer stimuli fra samme akustiske kontinuum – men ikke mindst om der er sammenfald mellem diskriminationstoppe og (eventuelt sprogspecifikke) kategorigrænser. Hvis den kategoriske effekt skyldes en psykoakustisk tærskel (jf. side 43ff), kan man ikke alene forvente, at den determinerer fonetiske kategorigrænser, men også diskriminationstoppene – uanset om den svarer til de lingvistiske grænser eller ej (Repp 1984). Hvis den kategoriske effekt derimod skyldes fonetisk kategorisering, må den forventes at forekomme ved alle lingvistiske kategorigrænser – og kun dér (Ades 1977).

Stimulus-kriterier

Udarbejdelsen af stimuli har naturligvis været styret af en vægtning mellem problemstillinger, jeg gerne ville undersøge og praktiske (tidsmæssige, tekniske og ressourcemæssige) begrænsninger. I et kontrolleret forsøg, som det jeg her har beskæftiget mig med, er stimulienes kvalitet af afgørende betydning – først og fremmest for resultaternes troværdighed. Derfor udgør udarbejdelsen af stimuli en overordentlig tidskrævende faktor. Som en konsekvens af dette har jeg opstillet en række betingelser, alle stimuli skal kunne opfylde.

1) For det første skal de anvendte kontinua så vidt muligt kun have ét varierende akustisk karakteristikum[3]. Dette krav har jeg naturligvis opstillet af hensyn til tolkninger og sammenligninger af forsøgsresultater.

2) For det andet skal stimuli virke så naturlige som muligt. Der er endnu ikke udført systematiske undersøgelser af denne faktors betydning for forsøg med KP. Jeg formoder dog, at man ville kunne spore en ringere kategorisk perception i forsøg med stimuli, der opfattes som unaturlige sproglyde, end med naturligt producerede stimuli (Liberman, Harris, Eimas, Lisker & Bastian 1961). Begrundelsen for denne antagelse er, at stimuli, der ikke opfattes som naturlige sproglyde, muligvis forleder forsøgspersonerne til

[3] Set ud fra et metodisk perspektiv er det en nødvendighed, at den afhængige variabel er af en helt anden karakter end den uafhængige variabel (dvs. den variabel, hvis ændringer man iagttager (Petersen 1995), hvis enhver form for cirkularitet skal undgås. Da formålet med dette forsøg er et studium af sprogperception, er det ganske naturligt at holde akustiske karakteristika som den kontrollerede (eller afhængige) variabel.

udelukkende at anvende auditive strategier, hvorimod naturlige sproglyde sandsynligvis begunstiger en fonetisk strategi (se Repp 1984 for en diskussion om strategier i forsøg med KP).

3) Endelig skal alle stimuli fra samme akustiske kontinuum have samme morfologiske status, dvs. at alle stimuli enten skal kunne opfattes som betydningsbærende tegn eller som betydningsløse sproglyde, men samme kontinuum må ikke indeholde stimuli fra begge kategorier. For at undgå en ukontrolleret semantisk variabel vil jeg endvidere tilstræbe, at stimuli, der kan opfattes som betydningsbærende tegn, har samme leksikalske status og anvendelsesfrekvens.

Valg af stimuli: et konsonantkontinuum

Til forsøget med KP af et konsonantkontinuum har jeg valgt at anvende et [t]-[d]-kontinuum med varierende VOT. Den primære årsag til, at valget faldt på netop denne fonetiske kontrast, er, at det med [t]-[d] er relativ let at skabe et akustisk kontinuum, der tager udgangspunkt i en naturligt produceret sproglyd: det er tilstrækkeligt gradvis at reducere aspirationslængden ved et naturligt produceret [t] for at skabe et akustisk kontinuum, der dækker en overgang fra [t] til [d].

[t] og [d] er to af de seks danske lukkelyde [p], [t], [k], [b], [d] og [g]. Fonemisk adskilles klusilerne p:b, t:d, k:g kun før (lige før eller som første element i en konsonantgruppe før) fuldvokaler i danske ord (jf. Basbøll 1997, 1999). I alle andre positioner sker der en systematisk neutralisation, dvs. ophævelse, af den forskel, der eksisterer i stavelsesinitial position før fuldvokal.

Den artikulatoriske fællesnævner for [t] og [d] er, at de begge dannes ved et alveolært lukke med tungespids (og -blad) mod alveolarranden (lige bag fortænderne i overmunden) og tungens sider mod kindtænderne i overmunden. Der er kun ringe variation mht. artikulationssted: I nogle tilfælde udgør lukket i [t] og [d] dog ikke et alveolært lukke, men et dentalt lukke. Palatogrammer viser oftest lige stor berøringsflade for [t] og [d] – når dette ikke er tilfældet, findes en anelse kraftigere ganeaftryk ved [d] (Hjelmslev 1954).

Den største forskel mellem [t] og [d] i initial position vedrører artikulationsmåden: [t] er stærkt aspireret, mens [d] er uaspireret eller meget lidt aspireret (målt vha. et aerometer, Fischer-Jørgensen 1966) og [t] produceres med åben glottis både under lukke og aspiration, hvorimod [d] produceres med lukket eller ganske lidt åben glottis (konstateret ved endoskopiske undersøgelser). Varigheden af [t]'s aspiration er meget varierende – mere end på f.eks. engelsk og tysk. Men netop i [t]'s aspiration ligger den væsentligste, muligvis eneste forskel fra [d], der i dansk – til forskel fra mange andre sprog – manifesteres ustemt[4]. Ingen hidtidige undersøgelser viser stærkere muskelspænding (tenuis) ved artikulationen af [t] end [d] – noget tyder på, at det måske forholder sig lige omvendt (altså at [t] er media)[5]. Eli Fischer-Jørgensens (1954, 1966) instrumentalfonetiske undersøgelser viser, at [d] altid har en kraftigere eksplosion end [t]. Desuden synes der at være den forskel, at eksplosionen i [t] slippes fri vha. luftstrømmen, hvorimod den i [d] frigøres aktivt ved bevægelse af tungespidsen. Denne beskrivelse understøttes endvidere af det faktum, at tungepresset er større ved [t] end [d]. I et komparativt perspektiv er det interessant, at [t] bedst kan karakteriseres ved en stærk aspiration kombineret med svag artikulation.

Kombinationen af stærk aspiration og svag artikulation er medvirkende til det forhold, at det danske [t] ikke alene er karakteriseret ved aspiration, men også affrikation. Affrikationen opstår som resultat af, at tungen slipper lukket langsomt, så der kommer en lille hæmmelyd inden pustet. Faktisk findes [t]'s affrikationsstøj, der følger lige efter eksplosionen, i akkurat samme frekvensområde som det danske [s] i f.eks. [asa] (Fischer-Jørgensen 1954, 1981). Ser vi på den akustiske intensitet, er det tydeligt, at aspirations- og affrikationsstøjen i [t] er kraftigere end den smule aspiration, der kan findes i [d].

[4] Denne beskrivelse kan dog i bedste fald kun dække, hvad sen-jespersenianerne betegner 'rigsmålsudtalen', hvilket er "den udtale der kan findes hos mennesker fra alle egne af landet og som ikke er hyppigere hos mennesker med lav social status end mennesker med høj social status" (Heger 1974: 45). F.eks. gælder det for bornholmsk, at [d] er stemt, mens [t] er en uaspireret, spændt lukkelyd (Hjelmslev 1954).

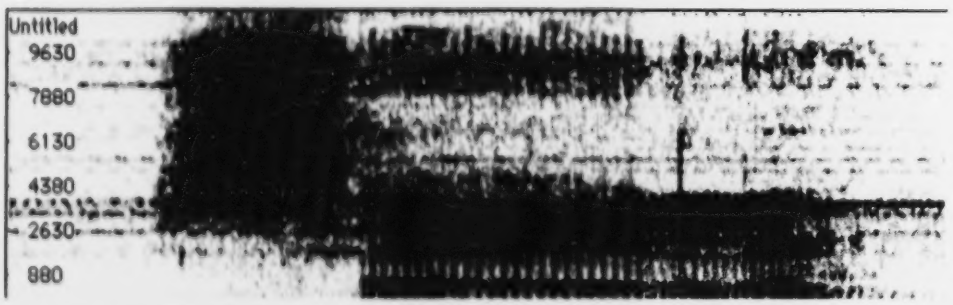
[5] Med dikotomien tenuis/media refererer jeg i dette tilfælde udelukkende til hhv. stærk og svag muskelspænding. Almindeligvis anvendes termerne tenuis/media som betegnelser for hhv. kraftig og svag eksplosion, men en kraftig eksplosion kan enten bero på stor muskelspænding eller kraftig luftstrøm, der igen kan skyldes kraftigt ekspirationstryk eller meget åben passage gennem stemmelæberne. Umiddelbart virker det måske mere hensigtsmæssigt at anvende termerne fortis/lenis som betegnelser for stærk og svag muskelspænding (jf. Hjelmslev 1954), men da man ved hæmmelyde ofte bruger dem specifikt om kombinationen af muskelspænding og ekspirationstryk, kan det blot skabe unødig forvirring at give dem en anden anvendelse i beskrivelsen af lukkelyde (Fischer-Jørgensen 1962).

Derimod tyder det på, at eksplosionen i [t] er kortere og har lavere intensitet. Eksplosionen i [t] kan – dog med lidt god vilje – skelnes fra den efterfølgende kraftige affrikationsstøj i målinger af støjens frekvensområde, jf. figur 2.2.1.

Ud fra et auditivt perspektiv er den eneste afgørende forskel mellem [t] og [d] aspirationen. Forsøg med ombytning af eksplosionsfasen i [t] og [d] har vist, at selve eksplosionen ikke har nogen betydning for perceptionen. Endvidere er en pause mellem eksplosionen og vokalen heller ikke tilstrækkelig for perceptionen af et [t]; tilstedeværelsen af aspirationsstøj er simpelthen nødvendig for en identifikation som [t] (Fischer-Jørgensen 1966).

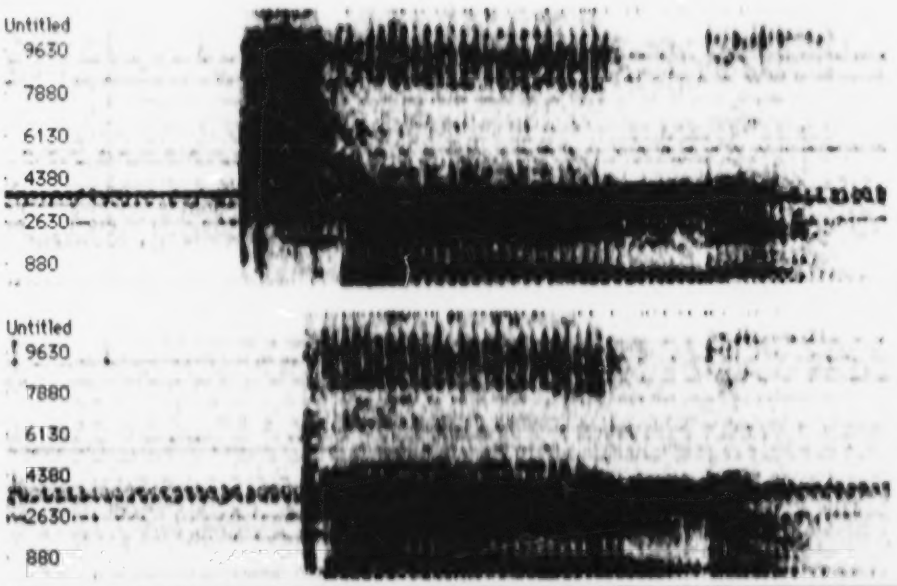
I løbet af årene 1994-98 har jeg udført en lang række pilotforsøg med [t]-[d]-stimuli (jf. f.eks. Jensen & Madsen 1995). Bl.a. har jeg forsøgt at øge (aspirations- og) eksplosionsintensiteten i et naturligt produceret [d] uden at det førte til en identifikation af stimulus som [t]. En svækkelse af eksplosions- og aspirationsintensiteten i et naturligt produceret [t] gav samme negative udfald[6]. Derimod har forsøg med gradvis reduktion af aspirationsstøj i et naturligt produceret [t] vist sig som en yderst pålidelig teknik til dannelsen af et akustisk kontinuum fra [t] til [d]. Jeg er ikke bekendt med nogen ulemper ved denne teknik.

Som det fremgår af afsnittet om motorhypotesen (side 49f), er det min hensigt at afprøve hypotesen om direkte korrespondens mellem sprog-perception og -produktion. Til dette formål skal jeg bruge en serie stimuli, der udgør et akustisk kontinuum, men som samtidig er mere eller mindre artikulatorisk kontinuerligt. I udarbejdelsen af det akustiske konsonant-kontinuum har jeg valgt at benytte mig af det forhold, at friktionsstøjen i et initialt [s] meget nøje svarer til aspirationsstøjen i et initialt [t].



[6] Medmindre teksten angiver andet, vil der herefter ikke blive skelnet mellem eksplosion og aspiration i behandlingen af danske lukkelyde.

BEST COPY AVAILABLE



Figur 2.2.1 Spektrogrammer af naturligt produceret 'sale', 'tale' og 'dale'. Bemærk den store lighed: spektrogrammerne for 'tale' og 'dale' er næsten identiske med spektrogrammet for 'sale'; de indeholder blot mindre akustisk energi før VO.

Der er dog to afgørende forskelle mellem initialt [s] og [t]: friktionsstøjens begyndelse og varighed. For det første er [s] en hæmmelyd og ikke en lukkelyd, dvs. at friktionsstøjen bliver gradvist kraftigere ved en stadig øget luftgennemstrømning. Friktionsstøjen i [t] starter derimod fra et lukke med en mindre eksplosion og virker derved som en abrupt overgang. Det er dog muligt at skabe samme abrupte overgang i det, der oprindeligt blev opfattet som et initialt [s]. Hvis man fjerner den første del af hvislelyden, så friktionen begynder meget brat, lyder den præcis som en eksplosion. Har man kun fjernet en meget lille del af friktionsstøjen i initialt [s], så VOT er meget højere end for et normalt produceret [t], vil den manipulerede sproglyd blive opfattet som et [ts]. De målinger, jeg har foretaget af initialt [t] – produceret af mig selv[7] – viser alle en VOT omkring 70 ms[8]. Reduceres VOT yderligere, vil den manipulerede sproglyd endelig blive opfattet som et [d].

[7] Jeg kan nok bedst karakteriseres som en yngre rigsmåls-/regionalsprogstalende fynbo.

[8] I et forsøg med dikotisk lytning i 1995 anvendte Jørgen Nørby Jensen og jeg en stimulus, der blev identificeret som initialt [t] (under 'forced choice'). Denne stimulus havde VO efter 80 ms. Eli Fischer-Jørgensens (1954) instrumentalfonetiske undersøgelse af danske lukkelyde indbefatter af uvisse årsager ikke initialt [t]'s VOT (men VOT for [b], [d], [g], [p] og [k]). Hun pointerer dog den ekstreme variation i aspirationslængde.

BEST COPY AVAILABLE

Som udgangspunkt for et VOT-kontinuum, der strækker sig fra [s] til [t] til [d], har jeg valgt et naturligt produceret [s]. For samtidig at opfylde de tidligere omtalte krav om naturlighed og ensartet morfologisk status vil stimuli blive udarbejdet som en overgang fra 'sale' (pl. af substantivet 'en sal') til 'tale' (infinitiven 'at tale' eller sg. af substantivet 'en tale') til 'dale' (enten pl. af substantivet 'en dal' eller infinitiven 'at dale'). Den bisyllabiske struktur 'CV-Cə' er samtidig den mest udbredte stavelsesstruktur i danske ord. Naturligvis kunne jeg have valgt andre ord end lige 'sale', 'tale' og 'dale' – også monosyllabiske – blot de kunne udgøre minimale par med [s], [t] og [d] – f.eks. bogstavnavnene for 'c', 't' og 'd'. Der er dog visse fordele ved 'sale', 'tale' og 'dale', der gjorde, at valget faldt på netop dem. Et væsentligt kriterium var, at jeg ønskede at anvende en akustisk stabil fuldvokal, dvs. en så lang 'steady state' vokallyd som muligt. Et af de væsentlige akustiske træk for perceptionen af lukkelyde er F2-transitionerne, der signalerer artikulationssted. Som beskrevet i Delattre et al. (1955) og Stevens et al. (1956) er resonans-frekvensen ved et alveolært lukke omkring 1800 Hz – stort set uafhængigt af den efterfølgende vokal (jf. afsnittet om 'locus' side 13ff). Dette betyder, at en efterfølgende vokal med F2 i et væsentlig lavere eller højere frekvensområde ville fremtvinge transitioner i F2. Det fremgår af figur 2.2.1, at jeg har undgået det problem. Hvis jeg havde valgt en anden vokal end [a], der netop på dansk har F2 omkring 1800 Hz, ville det betyde, at jeg ikke kunne hindre en manipulation af den betydningsfulde F2-transition. Valget af [a] var altså motiveret af kravene om kun ét varierende akustisk karakteristikum og størst mulig naturlighed.

Udarbejdelse af VOT-kontinuum

Som første trin i udarbejdelsen af et VOT-kontinuum indtalte jeg ordet 'sale' på en bærbar pc ved hjælp af dens indbyggede mikrofon. 'Sale' blev indtalt direkte på en 16-bit lydfil, der dækker området fra 0 til 44.100 Hz.

Udtalen kan bedst karakteriseres som 'afslappet udtale af den isolerede form', uden schwa-assimilation og i overensstemmelse med den form, der angives i Peter Molbæk Hansens *Dansk udtale* (1990) og Lars Brink et al.'s *Den store danske udtaleordbog* (1991).

Dernæst gik jeg i gang med at tilpasse den alt for lange lydfil til det digitaliserede signal. Resultatet blev en 709,32 ms lydfil (±0,0227 ms[9]), jf. figur 2.2.2.

Fillængde	709,32 ms
Signallængde	551,86 ms
Friktionslængde	125,00 ms
Vokallængde	584,32 ms

Figur 2.2.2

En række kardinalpunkter kan både måles i relation til begyndelsen af lydfilen og begyndelsen af signalet (herefter vil alle målinger relateres til begyndelsen af signalet):

	fil-start:	**signal-start:**
Signal-start	100,00 ms	**0 ms**
Signal-slut	651,86 ms	**551,86 ms**
VOT	225,00 ms	**125,00 ms**

Figur 2.2.3

Det eneste, der nu krævedes for at skabe et akustisk kontinuum, var en trinvis reduktion af VOT. Jeg valgte at holde alle faktorer konstante bortset fra, at jeg for hvert trin erstattede de første 10 ms friktionsstøj med stilhed. Figur 2.2.4 nedenfor angiver de 14 stimulis VOT[10]:

Stimulus	**1**	**2**	**3**	**4**	**5**	**6**	**7**
VOT (ms)	125	115	105	95	85	75	65

Stimulus	**8**	**9**	**10**	**11**	**12**	**13**	**14**
VOT (ms)	55	45	35	25	15	5	-5

Figur 2.2.4

Hvis man sammenligner det akustiske kontinuum i figur 2.2.4 med mine egne målinger af VOT ved initialt [t], der oftest har VO efter 70-75 ms, kan man

[9] Alle tidsmålinger er behæftet med denne ekstremt lille usikkerhed. Da alle stimuli blev indtalt på lydfiler, der rummer 44.100 Hz., havde de mindste dele jeg kunne manipulere, en længde på (1 sekund ÷ 44.100 Hz. = 0,0000227 sekund.) 0,0227 ms.
[10] De i alt 14 lydfiler betegnes C-01.wav - C-14.wav.

forvente, at i hvert fald stimulus 6 og 7 vil tilhøre den fonetiske [t]-kategori. Desuden er det værd at bemærke, at jeg med stimulus 14 har valgt at reducere signalet selv efter VO. Formålet med dannelsen af den – på dansk unaturlige – stimulus er at undersøge, om den akustiske stemthedsgrænse korresponderer med en kategorisk effekt. Det er min forventning, at den akustiske stemthedsgrænse vil markere en stigning i diskriminationen mellem stimulus 13 og 14 (i sammenligning med f.eks. stimulus 12 og 13).

Tekniske data

Udarbejdelsen af forsøget med KP lader sig dele op i flere faser. Der har således været anvendt forskelligt udstyr til indtaling og manipulation af stimuli, selve præsentationsprogrammet samt til behandlingen af de endelige forsøgsresultater.

Til udarbejdelsen af et VOT-kontinuum anvendte jeg en bærbar pc, en Toshiba 220CS med indbygget 16-bit lydkort (venligst udlånt af IT-Funktionen, Humaniora, Odense Universitet). Den første stimulus blev vha. den indbyggede mikrofon indtalt direkte på en to-spors 16-bit lydfil (*.wav-fil af PCM-typen, 44.100 Hz) i lydbehandlingsprogrammet GoldWave ver. 2.0 (copyright Chris S. Craig 1993).

GoldWave-programmet gør det muligt at indspille og afspille forskellige lyde på wav-filer af varierende længde. Der kan endvidere opstilles separate oscillogrammer (men ikke spektrogrammer) for lydene i hhv. venstre og højre kanal. En lang række funktioner muliggør desuden forskellige former for manipulation med de indspillede lyde, bl.a. ændre amplitude, indsætte pause, etc. GoldWave-programmet blev valgt, fordi det – i modsætning til de fleste andre lydbehandlingsprogrammer – gør det muligt at manipulere separat med lydene i de to kanaler, en funktion, der vil blive beskrevet nærmere i forbindelse med forsøg med dikotisk præsentation (side 121f).

Valg af stimuli: et vokalkontinuum

Til forsøget med KP af et vokalkontinuum har jeg valgt at anvende et [i]-[e]-[æ]-[a]-kontinuum, der varierer mht. F1 og F2. Til forskel fra det akustiske VOT-kontinuum, der repræsenterer konsonantserien, var det praktisk talt umuligt for mig at danne et vokalkontinuum, der udspringer af naturligt producerede sproglyde. Selvom vokalerne i højere grad end konsonanterne udgør et artikulatorisk kontinuum, kan vi ikke udtale dem, så de manifesteres

som akustisk ekvidifferente vokallyde. Jeg var med andre ord nødsaget til at anvende syntetisk producerede vokallyde.

[i], [e], [æ] og [a] udgør 4 af de 10 danske vokalfonemer[11]. De er alle stemte lyde, dvs., at stemmelæberne er i vibration. Desuden er vokalerne karakteriseret ved at have svælg- og mundhulrummet som resonansrum; ved udtalen af danske vokaler er næsehulrummet lukket ved at ganesejlet er hævet og således blokerer for luftgennemstrømning[12]. Resonansrummet kan varieres på 3 måder: ved ændret læbestilling, artikulationssted og åbningsgrad. Læbestillingen er bestemmende for, om resonansrummets åbning er stor eller lille. Alene ved at ændre læbernes stilling, kan vi altså ændre vokalen. Desuden kan vi ændre resonansrummets form ved at ændre tungens position i munden. Og endelig ændrer vi resonansrummets volumen/rumfylde, når vi ændrer på mundens åbningsgrad.

[i], [e], [æ] og [a] er alle kendetegnet ved at være fortungevokaler, dvs. at de artikuleres med hele tungemassen skudt frem og fortungen hævet mod den hårde gane. Endvidere adskiller disse 4 fortungevokaler sig artikulatorisk fra de resterende fortungevokaler ved ikke at have læberunding. Der er dog forskel på de 4 vokalers åbningsgrad: [i] betegnes typisk som snæver, [e] som halvsnæver, [æ] som halvåben og [a] som åben[13]. Vi må dog ikke glemme, at der i almindeligt talesprog er tale om glidende overgange mellem f.eks. fortunge og bagtunge, åben og snæver etc. Således viser flere palatogramundersøgelser af de urundede fortungevokaler [i], [e], [æ] og [a] (se f.eks. Fischer-Jørgensen 1962), at artikulationssted korrelerer med åbningsgrad: jo større kæbesænkning, desto mere tilbagetrukket tungemasse[14].

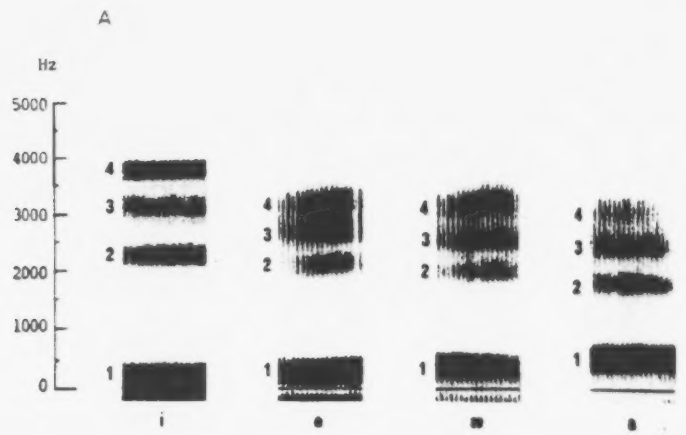
[11] Antallet af (vokal-)fonemer afhænger naturligvis af, hvordan man ønsker at fortolke en række manifestationer. Jeg vil ikke gå ind i en diskussion af det danske foneminventar her, da det ikke har (direkte) relevans for undersøgelsen eller anvendte stimuli.

[12] Jeg ser her bort fra dialektale træk som f.eks. de fynske nasalvokaler. Nasalvokaler er desuden almindelige i andre sprog, f.eks. polsk [ę] og [ą] i f.eks. 'przysięgnąć' (sværge) og de franske nasalvokaler i 'un bon vin blanc' (en god hvidvin).

[13] Jeg har valgt terminologien snæver, halvsnæver, halvåben og åben, bl.a. fordi den svarer til den engelske terminologi close, half-close, half-open og open. Otto Jespersen anvendte udtrykkene nær, halvnær, halvfjern og fjern, andre taler om højtstillet og lavtstillet, lukket og åben (se Heger 1981 for en diskussion om terminologien). Det er vigtigt at pointere, at jeg ved åbningsgrad fokuserer på tungemassens toppunkt i relation til ganen/ganesejlet og ikke berøringsfladen, som det er tilfældet hos Fischer-Jørgensen (1962).

[14] Denne relation er ikke et specifikt dansk fænomen, men skyldes et rent fysiologiske forhold, nemlig kæbens "hængsel".

Det væsentligste i en akustisk beskrivelse af vokalerne er, at de er rent periodiske, dvs. at de ikke indeholder komponenter, der står i uharmonisk forhold til grundtonen (herefter F0). Desuden er de karakteriseret ved en udpræget formantstruktur, dvs. intensiteten er koncentreret i visse frekvensområder (jf. beskrivelsen af formanter side 8ff). Tilsyneladende har F4 og højere formanter ikke nogen synderlig betydning for vokalernes forskellighed – de er nogenlunde ens for forskellige vokaler, men kan variere fra person til person[15]. F1 og F2 er derimod meget vigtige for alle vokaler. Det er først og fremmest deres position, der bestemmer vokalkvaliteten; de skifter fra vokal til vokal og har nogenlunde samme position for samme vokal udtalt af forskellige personer, jf. figur 2.2.5.



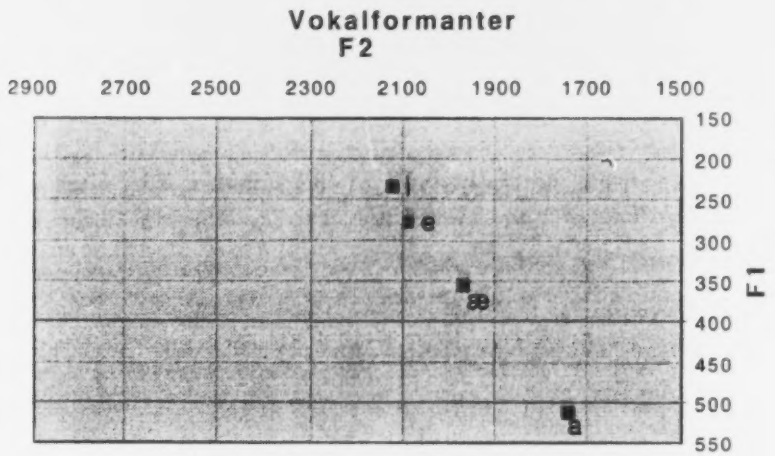
Figur 2.2.5 Spektrogrammer af danske lange vokaler – her de urundede fortungevokaler [i:], [e:], [æ:] og [a:]. Tallene 1-4 angiver formanter. (Fischer-Jørgensen 1981:22).

Når man taler om akustiske træk, der har afgørende betydning for perceptionen og identifikationen af vokaler, fokuseres der primært på de to første (relativt kraftigste) overtoneområder, F1 og F2. Hvis man opstiller de to første formanter i et todimensionalt diagram med F1 lodret og F2 vandret, får man en figur, der i forbløffende grad minder om en artikulatorisk

[15] F4 og højere formanter medvirker til dels til den individuelle stemmeklang, bl.a. gør kraftige høje formanter stemmen mere lys og skarp (Fischer-Jørgensen 1962). F3 er hos de fleste meget svag ved bagtungevokalerne og bidrager her næppe meget til vokalklangen, mens den er kraftigere ved fortungevokalerne og muligvis har en vis indvirkning på perceptionen og identifikationen (Thorsen & Thorsen 1976). Ofte præsenteres vokaler i et todimensionalt diagram, hvor den ene akse repræsenterer F2' (læs 'F2 mærke'), der er en "vægtet gennemsnitsformant" beregnet ud fra F2 og F3 (Fant 1958, Frøkjær-Jensen 1966).

BEST COPY AVAILABLE

vokalfigur (med åbningsgrad lodret og artikulationssted vandret), jf. figur 2.2.6.



Figur 2.2.6 Et todimensionalt diagram med F1 lodret og F2 vandret. De markerede områder repræsenterer de danske lange urundede fortungevokaler [i:], [e:], [æ:] og [a:]. Frekvensværdierne udgør et gennemsnit af formantfrekvenserne hos 10 mandlige forsøgspersoner under udtalen af langvokaler i en bisyllabisk 'CV:-Cə' -struktur. Diagrammet er baseret på data fra Frøkjær-Jensen (1966:35).

Figur 2.2.6 har en stor og ikke tilfældig lighed med artikulatoriske vokalfigurer, der illustrerer forholdet mellem åbningsgrad og artikulationssted. De urundede fortungevokaler adskiller sig fra hinanden ved åbningsgrad: [i:], [e:], [æ:] og [a:] er hhv. snæver, halvsnæver, halvåben og åben. Samtidig korresponderer åbningsgraden med artikulationsstedet således, at fortungevokaler udtales med mere og mere tilbagetrukket tungemasse, jo mere åben kæbestillingen er. Der er altså – ikke overraskende – en korrelation mellem taleorganernes position og formantstrukturen i artikulationens produkt, dvs. det akustiske signal (jf. side 9f).

Som en meget simpel tommelfingerregel for formantfrekvensernes beliggenhed kan man sige, at F1 hænger sammen med åbningsgrad, mens F2 vedrører artikulationssted og runding (Fischer-Jørgensen 1962, Thorsen & Thorsen 1976, Frøkjær-Jensen 1966). På trods af at relationerne mellem resonansrum og formanter ikke er helt simple, er det i dag muligt at beregne formanterne ret nøjagtigt, hvis man blot kender læbeåbningen, stedet for den

BEST COPY AVAILABLE

stærkeste indsnævring samt tværsnitsarealet på dette sted. Når man husker på, at store hulrum har dyb resonans og små hulrum høj resonans, er det muligt at se sammenhængen mellem artikulationen og formantstrukturen i [i], [e], [æ] og [a]. Ved de mest åbne fortungevokaler er resonanshulrummet bag tungens toppunkt mindre end ved de snævre, hvilket resulterer i en højere F1. Som tidligere omtalt medfører en stigning i åbningsgrad, at tungemassen trækkes tilbage og derved skaber et større resonanshulrum foran tungens toppunkt, hvilket manifesteres ved et fald i F2. En gradvis forøgelse af fortungevokalernes åbningsgrad kan altså med al rimelighed fortolkes som en bevægelse fra snævre fortungevokaler mod åbne bagtungevokaler[16].

Resultatet af ovenstående analyser og overvejelser var, at jeg var nødt til at danne et akustisk vokalkontinuum med to varierende faktorer, nemlig F1 og F2, hvis stimuli skulle svare til formantstrukturen i naturligt producerede vokallyde. I første omgang valgte jeg – ved hjælp af en synthesizer (se næste afsnit) – at danne en række stimuli baseret på analyser af min egen vokalproduktion. Desværre viste resultatet sig ikke at være tilfredsstillende – de syntetiske stimuli lød ikke tilnærmelsesvis som naturligt producerede vokallyde, hvorfor jeg måtte opgive denne strategi. Årsagen til at de syntetiserede lyde havde så ringe auditiv lighed med naturlige vokallyde er muligvis, at den synthesizer, jeg anvendte, udelukkende kunne variere F1 og F2, men med en prædefineret F3 på ca. 3500 Hz og F4 på ca. 3900 Hz. Til sammenligning kan det oplyses, at F3 ved udtalen af [i] og [a] hos mænd er omkring 3000/2300 Hz, hos kvinder 3400/2900 Hz og hos børn 3500/3050 Hz[17].

Da jeg ikke kunne danne et syntetisk vokalkontinuum baseret på min egen vokalproduktion, var jeg nødsaget til at anvende en auditiv metode. I samarbejde med Kirsten Gregersen fandt jeg vha. talesynthesizeren, hvad jeg rent auditivt opfattede som bedste eksempel på et [i] og et [a]. Disse to skulle udgøre endepunkterne for en lineær funktion mellem F1 og F2 – altså det akustiske vokalkontinuum, jf. figur 2.2.7.

[16] Denne beskrivelse er helt i overensstemmelse med flere nyere generelle vokalmodeller, bl.a. Basbøll & Wagners (1985) vokalmodel, der ikke indbefatter et hjørne for åbne fortungevokaler.

[17] Disse data stammer fra Frøkjær-Jensens (1966) analyser af 25 danskeres vokalproduktion. Forskellene mellem mænd, kvinder og børns formantfrekvenser forårsages af fysiologiske forskelle: kvinder og børn har et mindre resonanshulrum, først og fremmest et kortere svælg; i gennemsnit ligger formantfrekvenserne hos kvinder 17% højere end hos mænd.

	[i]	[a]
F1	200	500
F2	2800	2000

Figur 2.2.7

Jeg valgte at dele det beregnede akustiske kontinuum op i 18 lige store intervaller, hvilket resulterede i 19 stimuli. Valget af netop 19 formantpar som stimuli skyldes en række overvejelser om forholdet mellem antallet af stimuli og de sprogspecifikke kategorier.

For det første skal der være grundlag for en sammenligning mellem de to kontinua. Da det udarbejdede konsonantkontinuum udgjorde 12 stimuli og forventedes at dække 3 fonetiske kategorier, skal det syntetiske vokal-kontinuum med forventede 4 kategorier indeholde flere end de 12 stimuli.

For det andet skal der være grundlag for en sammenligning af kategoriseringer ved præsentation af samme akustiske kontinuum for forsøgs-personer med forskellige modersmål. Hvis jeg havde valgt at inddele det syntetiske vokalkontinuum i relativt få stimuli, ville der være en vis risiko for sammenfald mellem forskellige sprogs kategorier – et sammenfald udelukkende determineret af alt for store akustiske spring.

Omvendt er der også en ulempe forbundet med en så fin opdeling af det syntetiske kontinuum. Det er ganske klart, at en ringe diskrimination mellem meget forskellige stimuli er et langt vægtigere argument for kategorisk perception end ringe diskrimination mellem stimuli, der kun varierer ad ganske små kvantitative skridt.

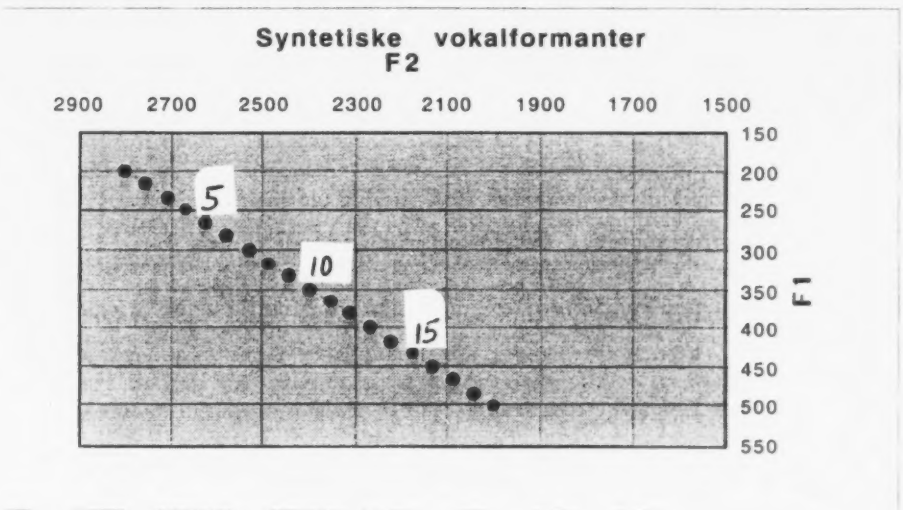
For at skabe et akustisk kontinuum bestående af 19 stimuli – fra [i] til [a], jf. figur 2.2.7 – fandt jeg differencen mellem hhv. [i] og [a]'s F1- og F2-frekvenser (F1: 300 Hz, F2: 800 Hz). Derefter dividerede jeg differencerne med 18, hvorved jeg fandt afstanden mellem hver stimulus; for hvert skridt fra [i] mod [a] stiger F1 med 16,66 Hz og F2 falder med 44,44 Hz. De to parametre F1 og F2 blev således inddelt i 19 punkter som illustreret i figur 2.2.8.

Stimulus	1	2	3	4	5	6	7	8	9	10
F1	200	217	233	250	267	283	300	317	333	350
F2	2800	2756	2707	2667	2622	2578	2533	2489	2444	2400

Stimulus	11	12	13	14	15	16	17	18	19
F1	367	383	400	417	433	450	467	483	500
F2	2356	2311	2267	2222	2178	2133	2089	2044	2000

Figur 2.2.8

I et F1-F2-diagram som figur 2.2.9 fremgår det tydeligt, at de 19 syntetiske stimuli udgør et vokalkontinuum – vel at mærke et akustisk kontinuum, der skal foregive at være et artikulatorisk kontinuum fra en snæver fortungevokal [i] til et åbent, tilbagetrukket [a].



Figur 2.2.9

Udarbejdelse af et formantkontinuum

Da jeg først havde beregnet det akustiske kontinuum, manglede jeg blot at realisere de syntetiske stimuli. Til udarbejdelsen af de konkrete stimuli anvendte jeg en talesynthesizer (jf. afsnittet med tekniske data, side 58), der kan variere F1, F2 samt F0 (grundtonen). For ikke at lade grundtonen have indflydelse på diskriminationen, valgte jeg at lade den være konstant omkring 140 Hz (se FFT-analyse af stimuli i bilag A).

Ved hjælp af en talesynthesizer, der var direkte forbundet med en PowerPC, kunne jeg nu foretage spektralanalyser af synthesizerens output. Proceduren var ganske simpel: først valgte jeg et koordinat på synthesizerens eksterne F1-F2-diagram, hvorefter jeg foretog en spektralanalyse af det akustiske signal. Naturligvis indebar denne procedure, at jeg gang på gang måtte justere koordinaterne, for at få dem til at svare til de prædefinerede formantværdier. Når jeg endelig havde fundet en koordinat, der kunne producere de ønskede formantfrekvenser, blev det akustiske signal gemt på en ét-spors 16-bit lydfil (*.sig-fil, 22050 Hz) vha. lydbehandlingsprogrammet Signalyze™.

I første omgang havde jeg gemt alle stimuli i lydfiler af 2000 ms varighed. –Men syntetisk producerede vokallyde af den længde minder slet ikke om naturligt producerede sproglyde. Næste opgave var derfor at reducere det akustiske signal til 350 ms. Tidligere undersøgelser af vokallængdens effekt viser entydigt, at lange vokallyde (200-300 ms) giver en relativt høj diskrimination, mens korte vokaler (50-150 ms) giver en lav diskrimination (Pisoni 1971, 1973, 1975; Sachs 1969). Jeg valgte her at lade en relativt langvarig vokal (350 ms) kompensere for den store sværhedsgrad, som blev fremkaldt af de meget små akustiske forskelle. Det skal dog samtidig tilføjes, at de syntetiske vokallyde herefter blev tilspidset over 50 ms i hver ende (tapering), hvilket resulterer i 250 ms konstant vokallyd[18].

Endelig blev samtlige 19 lydfiler konverteret fra *.sig til *.wav-filer ved hjælp af et konverteringsprogram, FileConverter 2.11.1 (udviklet af Stefan Werner, freeware), som er en del af Signalyze-pakken.

Min tidlige skepsis ved syntetisk producerede vokallyde blev afløst af en stor tilfredshed med det endelige resultat; ikke alene kan de syntetiske signaler med lethed identificeres som vokallyde, de lyder heller ikke så unaturligt, som jeg havde frygtet.

Tekniske data

Til udarbejdelsen af et syntetisk vokalkontinuum anvendte jeg en talesynthesizer, FONEMA OVE I resonator, med variabel F1 og F2, men

[18] 250 ms signal er dog stadig en relativ lang langvokal. Niels Reinholt Petersen (1976) har i et forsøg med KP af kort versus lang vokal fundet, at danske sprogbrugeres kategorigrænse er omkring 140 ms. Forsøgspersoner blev endvidere bedt om at tilpasse fuldvokalens længde i ordet 'læse' "...in accordance with a criterion of "naturalness"" (Petersen 1976: 69). Det viste sig, at forsøgspersonerne foretrak en varighed over eller omkring 200 ms for en langvokal.

invariabel F3 og F4 (se side 62). OVE I er i modsætning til de senere generationer OVE II og OVE IIIa-d kun i stand til at producere syntetiske vokallyde, ikke aperiodiske signaler. Vokalsynthesizeren var venligst udlånt af Fonetisk laboratorium, Odense Universitet. OVE I blev forbundet direkte med en PowerPC via linjeindgangen.

Den software, der blev anvendt til analyse og manipulation af stimuli, er et stort lydbehandlingsprogram, Signalyze™, som Fonetisk Laboratorium på Odense Universitet købte licens til for et par år siden. Signalyze™ version 3.12 (copyright Eric Keller 1988-1994) er et omfattende, kraftfuldt program, der primært er designet til analyse af sprog. Til analysen af de akustiske signaler (jf. akustiske analyser i bilag A) anvendte jeg et FFT-spektrogram (Fast Fourier Transform) med 'wide band' (8 ms/125 Hz) samt et FFT-spektrum (0-5512 Hz)[19]. Alle, der har prøvet at arbejde med akustisk analyse ved, at det er meget svært at foretage præcise målinger. På trods af stor grundighed er alle frekvensværdier da også behæftet med en usikkerhed på ± 5 Hz.

Procedure: præsentations- og registreringsprogram

Til forsøgene udviklede jeg et program specielt konstrueret til kontrolleret præsentation af akustiske stimuli og registrering af forsøgspersoners respons. Programmet fik navnet 'DikoTest 3.0'.

De to første versioner af DikoTest blev egentlig udviklet i 1994-95 i et samarbejde mellem Nikolaj Frandsen, Jørgen Nørby Jensen og undertegnede (Jensen & Madsen 1995). På daværende tidspunkt arbejdede Jørgen og jeg med dikotisk lytning (deraf navnet 'DikoTest'). Det var egentlig min intention at anvende samme program – dog i en modificeret form – til forsøg med KP, men da den oprindelige kode desværre er gået tabt, var jeg nødt til at programmere en helt ny version. DikoTest 3.0 har jeg udviklet ved hjælp af 'Microsoft Visual Basic 3.0', der er et komplekst og kraftfuldt program til udarbejdelse af softwareprogrammer til pc'er (kræver Windows som styresystem).

DikoTest 3.0 er konstrueret, så det kan afspille lydfiler, præsentere tekst på skærmen samt registrere forsøgspersoners respons på tastaturet. Hele denne

[19] For yderligere tekniske data se Keller 1994.

procedure varetages af hhv. en styrefil (*.csv-fil) og en registreringsfil (*.csr-fil).

Præsentationsserier og registrering

Styrefilerne består af præsentationsserier, der alle er bygget op efter samme princip: først defineres en *.wav-lydfil, som DikoTest kan afspille i medie-afspilleren. Dernæst præsenteres den tilhørende tekst på skærmen:

Identifikationsøvelser[20]:

Konsonanter:	'HØRER DU ET D, S ELLER T? SVAR MED ÉN TAST. #XX'
Vokaler:	'HØRER DU ET A, E, I ELLER Æ? SVAR MED ÉN TAST. #XX'

Diskriminationsøvelser (både konsonanter og vokaler):

Præsentation af lyd 1:	'LYD 1. #XX'
Præsentation af lyd 2:	'LYD 2. ENS ELLER FORSKELLIGE? #XX'

DikoTest 3.0 giver mulighed for automatisk præsentation med faste intervaller af 1 til 5000 ms varighed. Jeg havde dog valgt ikke at anvende den automatiske præsentation, men derimod lade forsøgspersonerne selv bestemme hastigheden; når forsøgspersonen havde afgivet respons på tastaturet (eller hørt lyd 1), skulle han/hun selv trykke på tasten [pil ned] for en ny præsentation.

Når forsøgspersonerne responderede på akustiske stimuli og tekst ved tryk på tastaturet, blev lydfilens navn, den præsenterede tekst, responsen samt reaktionstiden registreret i en registreringsfil (*.csr-fil). DikoTest 3.0 giver endvidere mulighed for indskrivning af oplysninger om forsøgspersonerne.

Som udgangspunkt for præsentationsseriernes opbygning blev der opstillet en bestemt rækkefølge af stimuli i regnearkprogrammet Microsoft Excel.

For det akustiske S-T-D-kontinuum blev der fremstillet en identifikationsserie bestående af 28 præsentationer, dvs. at hver af de 14 stimuli blev præsenteret 2 gange i en serie. Diskriminationsserien bestod derimod af 40 stimuluspar, der var kombineret, så hver stimulus blev kontrasteret med den foregående,

[20] Jeg valgte at opstille svarmulighederne i alfabetisk rækkefølge, da jeg forventer, at kun ganske få af forsøgspersonerne har viden om fonetiske systemer. '#xx' repræsenterer et præsentationsnummer, som forsøgspersonerne skulle notere, hvis deres respons via tastaturet var forkert, dvs. utilsigtet.

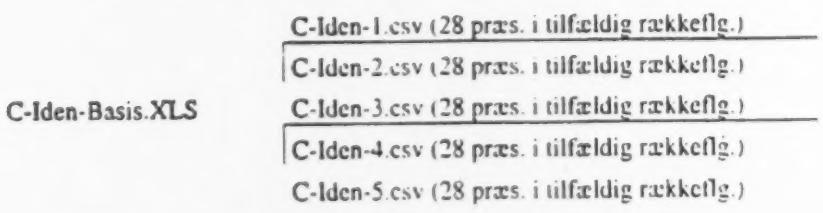
den samme og den efterfølgende stimulus i det akustiske kontinuum. F.eks. blev stimulus 3 (C-03.wav) kontrasteret med stimulus 2 (C-02.wav), stimulus 3 og stimulus 4 (C-04.wav), og stimulus 4 med 3, 4 og 5, etc[21]. Således bestod diskriminationsserier med konsonant-stimuli af 40 par ((14 · 3) - 2), hvoraf 14 var ens og 26 forskellige.

For det syntetiske vokalkontinuum blev der fremstillet en identifikationsserie bestående af 19 præsentationer, dvs., at hver af de 19 stimuli blev præsenteret 1 gang i en serie. Endelig bestod diskriminationsserien med vokalstimuli af 55 stimuluspar, der var kombineret efter samme princip som beskrevet ovenfor ved konsonantdiskrimination, dvs. 55 par ((19 · 3) - 2), hvoraf 19 var ens og 36 forskellige.

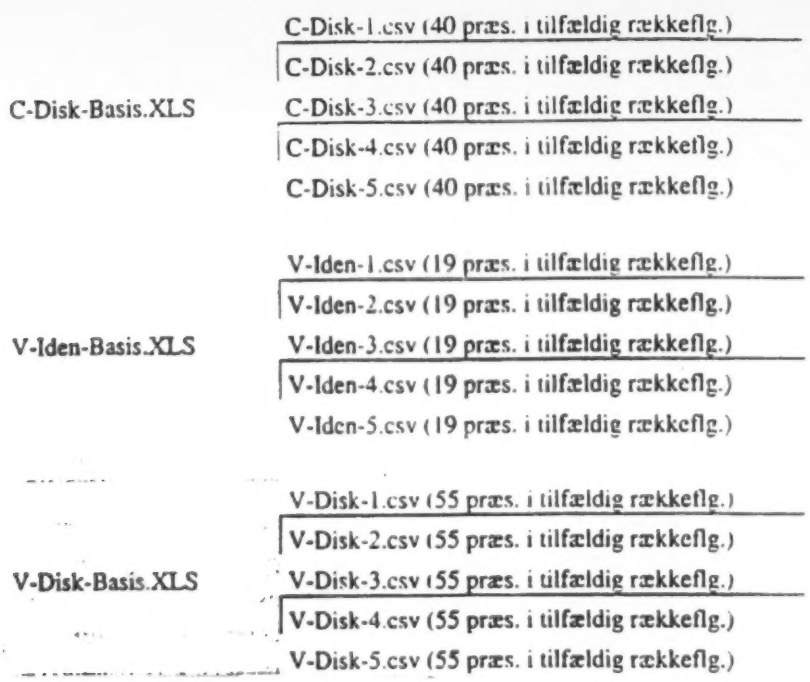
Endelig skal det tilføjes, at alle præsentationsserier blev indledt af en "introduktionsrunde" bestående af 6 identifikations- eller diskriminations-øvelser. Formålet med "introduktionsrunden" var udelukkende at give forsøgspersonerne mulighed for vænne sig til udstyr, procedure etc. Data fra disse "introduktionsrunder" er ikke medregnet som forsøgsresultater.

De 4 basisserier for konsonantidentifikation, konsonantdiskrimination, vokal-identifikation og vokaldiskrimination bliver herefter betegnet 'C-Iden-Basis.XLS', 'C-Disk-Basis.XLS', 'V-Iden-Basis.XLS' og 'V-Disk-Basis.XLS'

Præsentationsseriernes opbygning i de 4 basisserier er udgangspunkt for udviklingen af alle andre serier. Ved hjælp af funktionen 'slump' (i regneark-programmet Microsoft Excel) kunne jeg tilskrive hver præsentations-kommando en vilkårlig talværdi og derefter sortere dem med faldende værdier. Således genererede jeg for hver basisserie 5 styrefiler med hhv. 28, 40, 19 og 55 præsentationer i tilfældig rækkefølge, se figur 2.2.10.



[21] Jeg har valgt at dobbeltrepræsentere stimuli i kontrastserierne (f.eks. stimulus 3 versus stimulus 4 og 4 versus 3) for at sikre resultaterne mod bias pga. præsentationsrækkefølgen som intervenerende variabel.



Figur 2.2.10

Ved behandlingen af de endelige forsøgsresultater blev regnearkprogrammet Microsoft Excel ligeledes benyttet. De (4 · 5) 20 forsøgsseriers tilhørende resultatfiler blev placeret i regnearket og ved hjælp af kommandoen 'sorter' blev data (lydfilens navn, den præsenterede tekst, respons på tastatur samt reaktionstid) herefter ført tilbage i basisseriens oprindelige rækkefølge.

Forsøgsforhold

DikoTest-programmet blev installeret i en bærbar pc (se afsnit om tekniske data side 58), som blev brugt til alle forsøgene. Et par højkvalitets stereo-hovedtelefoner af mærket Sennheiser HD 480 blev ligeledes benyttet under alle forsøgene.

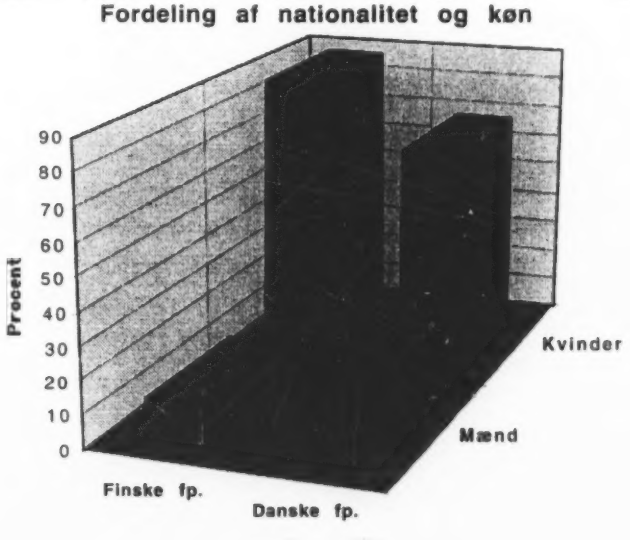
Alle forsøgene blev foretaget i perioden 12. juni - 10. juli 1998; forsøg med finske fp: 17. - 20. juni; forsøg med danske fp: 12. juni - 10. juli. Gruppen af finske forsøgspersoner deltog i juni måned i danskkurser (Sommer-universitet) på Odense Universitet, hvor de bl.a. blev undervist i dansk fonetik (af undertegnede). Undervisningen i dansk fonetik begrænsede sig til 6 dobbelttimer, men forventes dog at have haft en vis indflydelse på forsøgs-resultaterne. Forsøgene med finske studenter fandt sted på Dalum

Landbrugsskole og i et kontor på Scandinavian Area Studies, Odense Universitet. Gruppen af danske forsøgspersoner gennemgik oftest forsøgene i private hjem i hhv. Odense og Svendborg. Et forsøg tog i gennemsnit 25-35 minutter.

Forsøgspersoner

Alle forsøgspersonerne udfyldte et registreringsskema. De oplysninger, der har relevans for forsøg med KP, er først og fremmest forsøgspersonens nationalitet, modersmål, køn og alder[22].

For at deltage i forsøget skulle forsøgspersonerne være i besiddelse af en normal hørelse – dvs. de måtte ikke have fået konstateret en væsentlig nedsat hørelse eller andre høreskader.



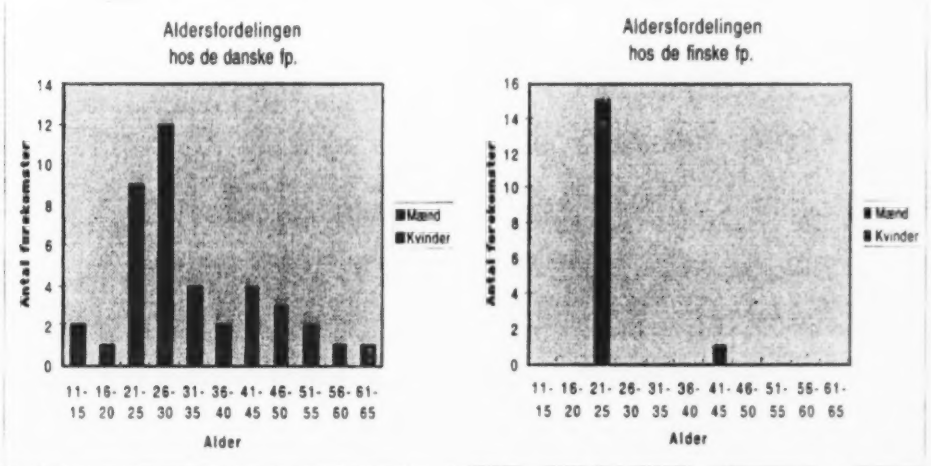
Figur 2.2.11

I forsøget deltog i alt 57 forsøgspersoner fordelt på 41 danskere og 16 finner, 42 kvinder og 15 mænd i alderen 12 - 61 år (gennemsnitsalderen var 28 år). Ingen af forsøgspersonerne medvirkede flere gange.

[22] Af registreringsskemaet fremgik det, at jeg udelukkende ville anvende forsøgspersonernes initialer. Desværre havde jeg ikke forudset, at der ville være en række sammenfald, hvorfor jeg måtte anvende forkortelser bestående af 3 bogstaver (oftest de to første bogstaver i fornavnet samt første bogstav i efternavnet).

Som det fremgår af figur 2.2.11, var der blandt de finske forsøgspersoner en stor overvægt af kvinder, hvorimod de danske forsøgspersoner var mere ligeligt fordelt over de to køn.

Også aldersmæssigt udgjorde de finske forsøgspersoner en konform gruppe: af de 16 finske deltagere var 15 (93,75 %) i alderen 21 - 25 år (gennemsnitsalderen var 23,6 år). Gruppen af danske forsøgspersoner var med en gennemsnitsalder på 32,5 år lidt ældre, men havde en langt større aldersmæssig spredning (12 - 61 år), jf. figur 2.2.12.



Figur 2.2.12

Det var ikke alle forsøgspersoner, der responderede på samtlige 4 testserier – enkelte deltagere sprang fra efter kun at have gennemgået én af de to hovedkategorier, konsonantidentifikation og -diskrimination eller vokalidentifikation og -diskrimination. Hvis en deltager f.eks. havde gennemgået præsentationsserie 'C-Iden-1', skulle vedkommende derefter gennemgå 'C-Disk-1' – og evt. 'V-Iden-1' og 'V-Disk-1'. Jeg forsøgte desuden at fordele det samlede antal forsøgspersoner over de 5 varianter af hver præsentationsserie. Formålet var naturligvis at undgå bias determineret af rækkefølgen i præsentationsserierne, jf. figur 2.2.13.

Figur 2.2.13

Hvis man ser bort fra parametrene nationalitet, køn og alder, er de to grupper af forsøgspersoner sammenlignelige: de er blevet præsenteret for de samme akustiske stimuli under ensartede forhold.

Danske fp:

	Køn	Alder	Præs.serie	Konsonantøvelse	Vokaløvelse
AAO	♂	57	3	Ja	Ja
BDL	♂	45	4	Ja	Ja
BRB	♂	27	2	Ja	Ja
CHW	♀	29	5	Nej	Ja
COM	♀	23	1	Ja	Ja
DOB	♀	34	1	Ja	Ja
FLJ	♂	47	3	Ja	Ja
GIM	♀	29	4	Ja	Nej
GNE	♀	27	4	Ja	Ja
HAF	♀	46	4	Ja	Ja
HEP	♀	49	2	Ja	Ja
HES	♂	32	2	Ja	Ja
IKM	♀	61	5	Ja	Ja
IMN	♀	30	1	Ja	Ja
JAO	♀	25	1	Ja	Ja
JEK	♂	27	1	Ja	Ja
JNJ	♂	26	3	Ja	Ja

JOJ	♀	24	2	Ja	Ja
JTR	♂	35	2	Ja	Ja
KAM	♀	36	5	Ja	Ja
KAO	♀	52	4	Ja	Nej
KIG	♀	52	5	Ja	Ja
KRR	♀	24	3	Ja	Ja
LDL	♀	41	5	Ja	Ja
LER	♀	24	1	Ja	Ja
LIL	♀	14	1	Ja	Ja
LOK	♀	25	1	Ja	Ja
MAN	♀	20	2	Ja	Ja
MCK	♀	24	3	Ja	Ja
MDL	♀	12	2	Ja	Ja
MOH	♂	26	1	Ja	Ja
MOS	♂	42	5	Ja	Ja
OLG	♂	45	4	Ja	Ja
PES	♂	28	1	Ja	Ja
SAH	♀	24	4	Ja	Ja
SMF	♀	24	3	Ja	Ja
SOD	♀	29	5	Ja	Ja
SUM	♀	36	4	Ja	Ja
SØH	♂	26	3	Ja	Ja
TES	♀	31	3	Ja	Ja
TRT	♀	26	2	Ja	Ja

Finske fp:

	Køn	Alder	Præs.serie	Konsonantøvelse	Vokaløvelse
ANK	♀	24	4	Ja	Ja
ELK	♀	21	1	Ja	Ja
JAH	♂	25	3	Ja	Ja
JON	♀	23	3	Ja	Ja
KAS[23]	♀	21	2	Ja	Ja
KAU	♀	23	4	Ja	Ja
KAV	♂	44	2	Ja	Ja

[23] KAS er fra Estland og burde derfor behandles separat fra de finske forsøgspersoner, men da hendes forsøgsresultater stemmer helt overens med de finske, har jeg valgt at behandle de to grupper samlet.

LES	♀	21	1	Ja	Ja
LOP	♀	22	1	Ja	Ja
LOR	♀	22	5	Ja	Ja
MAL	♀	22	5	Ja	Ja
MEO	♀	23	5	Ja	Ja
MIK	♀	21	1	Ja	Ja
OUS	♀	23	4	Ja	Ja
REK	♀	21	1	Ja	Ja
SHA	♀	22	2	Ja	Ja

Figur 2.2.14

2.3 Resultater fra forsøg med kategorisk perception: et S-T-D-kontinuum

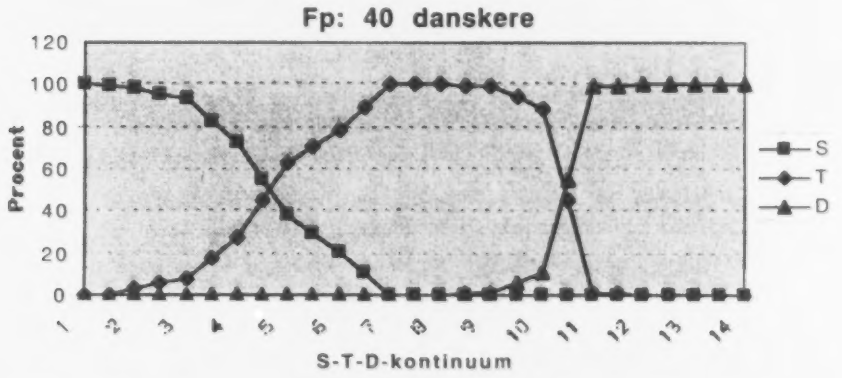
Det er min hensigt i det følgende at præsentere resultater fra forsøg med KP af et akustisk S-T-D-kontinuum, der udelukkende varierer mht. VOT. Resultaterne stammer fra to grupper forsøgspersoner, 40 danske og 16 finske.

Desuden vil jeg analysere de præsenterede data for at finde ud af, om de opfylder kriterierne for KP – og ikke mindst i hvilken grad. Senere vil jeg endvidere analysere resultaterne med henblik på de opstillede hypoteser i kapitel 2.2: Forsøg med kategorisk perception, nemlig motorhypotesen, hypotesen om sprogspecifikke kategorigrænser og hypotesen om stimulus-asymmetri.

Identifikation af stimuli fra et VOT-kontinuum

De første resultater, jeg vil præsentere, er fra de 4 identifikationsserier, hvor forsøgspersonerne 2 gange skulle identificere hver af de 14 stimuli som enten 'dale', 'sale' eller 'tale'.

Gruppen af danske fp bestod af 40 personer, der er blevet præsenteret for hver stimulus 2 gange, hvilket betyder, at procentfordelingen i figur 2.3.1 er baseret på 1.120 præsentationer og reaktioner (40 · 14 · 2).



Figur 2.3.1

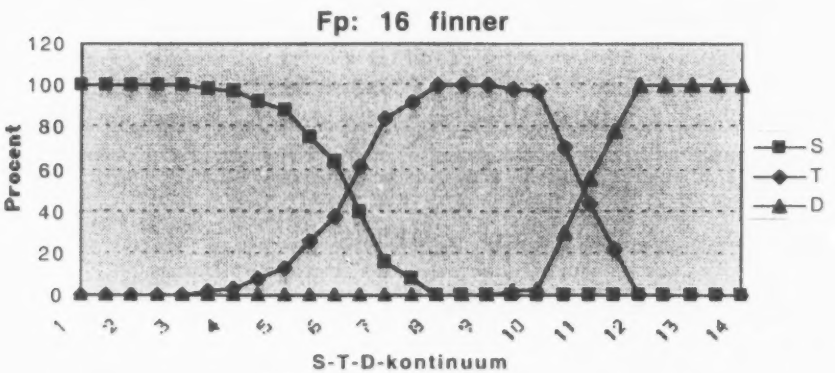
Det fremgår tydeligt af den trimodale (trepuklede) fordeling, at forsøgspersonerne har kunnet identificere alle stimuli som tilhørende én af de tre

BEST COPY AVAILABLE

sproglige kategorier, 'S, 'T eller 'D. Forsøgspersonerne har f.eks. ikke tilskrevet det akustiske kontinuum blot to kategorier. En platykurtisk fordelingsform – dvs. med "brede" kategoritoppe – kombineret med den høje identifikationsprocent er endvidere en god indikator for, at stimuli har virket relativt naturlige. Havde forsøgspersonerne oplevet mange stimuli som "unaturlige", f.eks. som en mellemting mellem to forskellige sproglyde, ville forskellen mellem de tre fordelingskurver være meget mindre, hvilket ville betyde en lavere og fladere fordeling omkring 50% percentilen.

I histogrammet over de finske forsøgspersoners identifikation af samme VOT-kontinuum, figur 2.3.2, finder vi samme høje identifikationsprocent og fordelingskurtose som i det danske.

Gruppen af finske fp bestod af 16 personer, der ligeledes blev præsenteret for hver stimulus 2 gange, hvilket gav i alt 448 præsentationer og reaktioner (16 · 14 · 2).



Figur 2.3.2

Ser vi nærmere på, hvordan de to grupper har tilskrevet kategorier til det akustiske kontinuum, træder nogle forskelle frem. Hvis vi udelukkende fokuserer på de observationsværdier, der overstiger 50% percentilen[1], kan vi se, at de danske forsøgspersoner har tilskrevet stimulus 1-4 til kategorien 'S, 5-10 til 'T og 11-14 til 'D, hvor de finske forsøgspersoner har tilskrevet stimulus 1-6 til 'S, 7-10 til 'T og 11-14 til 'D. De danske forsøgspersoner har

[1] 50% percentilen i figur 2.3.1 og 2.3.2 er ikke udtryk for medianen, da koordinaterne (i horisontalt perspektiv) ikke er normalfordelte (Petersen 1995).

altså en mindre S-kategori og en større T-kategori end de finske forsøgspersoner.

Statistiske analyser

Nu kunne det være godt at vide, om kategoritilskrivningen kan være opstået ved tilfældighed: alt andet lige er der jo 33% sandsynlighed for, at en stimulus tilskrives én af de 3 kategorier. Det er altså vigtigt, hvor stor sandsynligheden er for netop den konkrete kategoritilskrivning. For at besvare dette spørgsmål har jeg opstillet en såkaldt 'nul-hypotese', H_0, der skal afsløre sandsynligheden for, at en given stimulus ikke identificeres som tilhørende en given kategori på basis af de to populationer, danske identifikationer og finske identifikationer.

Sandsynlighed for H_0 for hver enkelt stimulus fremgår af P-værdierne i figur 2.3.3 Sandsynligheden er beregnet ud fra forholdet mellem observationsværdien (n), populationens aritmetiske middelværdi (X) og standard-afvigelse (S) set i forhold til en normalfordeling (den gaussiske fordeling). Sandsynligheden for H_0 er ved enkelte stimulus så lav, at den ikke indgår i min normalfordelingstabel (P som funktion af Z) (Petersen 1995). Disse sandsynligheder ($P < 0{,}001$) er i figur 2.3.3 markeret med tre stjerner (***). $H_0(P < 0{,}05)$ – dvs. en statistisk signifikant fordeling – er derimod markeret med én stjerne (*).

Population: danske forsøgspersoners identifikationer
Kategori: S

Stimulus	n	%	X	S	Z	P
1	80	100	1,0	0	0	***
2	78	98	0,975	0,157	-6,21	***
3	74	93	0,925	0,265	-3,49	***
4	58	73	0,725	0,449	1,62	0,053
5	30	38	0,375	0,487	-0,77	0,221
6	17	21	0,213	0,412	-0,52	0,302
7	0	0	0	0	0	1,0
8	0	0	0	0	0	1,0
9	0	0	0	0	0	1,0
10	0	0	0	0	0	1,0
11	0	0	0	0	0	1,0
12	0	0	0	0	0	1,0

BEST COPY AVAILABLE

13	0	0	0	0	0	1,0
14	0	0	0	0	0	1,0

Kategori: T

Stimulus	n	%	X	S	Z	P
1	0	0	0	0	0	1,0
2	2	3	0.025	0,157	-0.16	0,436
3	6	8	0.075	0,265	-0.28	0,39
4	22	28	0.275	0,449	-0.61	0,271
5	50	63	0.625	0,487	-1.28	0,1
6	63	79	0.788	0,412	-1.91	0,028*
7	80	100	1.0	0	0	***
8	80	100	1.0	0	0	***
9	79	99	0.988	0,112	-8.82	***
10	71	89	0.888	0,318	-2.79	0,03*
11	1	1	0.013	0,112	-0.12	0,452
12	0	0	0	0	0	1,0
13	0	0	0	0	0	1,0
14	0	0	0	0	0	1,0

Kategori: D

Stimulus	n	%	X	S	Z	P
1	0	0	0	0	0	1,0
2	0	0	0	0	0	1,0
3	0	0	0	0	0	1,0
4	0	0	0	0	0	1,0
5	0	0	0	0	0	1,0
6	0	0	0	0	0	1,0
7	0	0	0	0	0	1,0
8	0	0	0	0	0	1,0
9	1	1	0.025	0,112	-0.22	0,413
10	9	11	0.113	0,318	-0.36	0,359
11	79	99	0.988	0,112	-8.82	***
12	80	100	1.0	0	0	***
13	80	100	1.0	0	0	***
14	80	100	1.0	0	0	***

BEST COPY AVAILABLE

Population: finske forsøgspersoners identifikationer

Kategori: S

Stimulus	n	%	X	S	Z	P
1	32	100	1.0	0	0	***
2	32	100	1.0	0	0	***
3	32	100	1.0	0	0	***
4	31	97	0.969	0,177	-5.47	***
5	28	88	0.875	0,336	-2.60	0,004*
6	20	63	0.625	0,492	-1.27	0,102
7	5	16	0.156	0,369	-0.42	0,337
8	0	0	0	0	0	1,0
9	0	0	0	0	0	1,0
10	0	0	0	0	0	1,0
11	0	0	0	0	0	1,0
12	0	0	0	0	0	1,0
13	0	0	0	0	0	1,0
14	0	0	0	0	0	1,0

Kategori: T

Stimulus	n	%	X	S	Z	P
1	0	0	0	0	0	1,0
2	0	0	0	0	0	1,0
3	0	0	0	0	0	1,0
4	1	3	0.031	0,177	-0.18	0,429
5	4	13	0.125	0,336	-0.37	0,356
6	12	38	0.375	0,492	-0.76	0,224
7	27	84	0.844	0,369	2.29	0,011*
8	32	100	1.0	0	0	***
9	32	100	1.0	0	0	***
10	31	97	0.969	0,177	-5.47	***
11	14	44	0.438	0,504	-0.87	0,197
12	0	0	0	0	0	1,0
13	0	0	0	0	0	1,0
14	0	0	0	0	0	1,0

BEST COPY AVAILABLE

Kategori: D

Stimulus	n	%	X	S	Z	P
1	0	0	0	0	0	1,0
2	0	0	0	0	0	1,0
3	0	0	0	0	0	1,0
4	0	0	0	0	0	1,0
5	0	0	0	0	0	1,0
6	0	0	0	0	0	1,0
7	0	0	0	0	0	1,0
8	0	0	0	0	0	1,0
9	0	0	0	0	0	1,0
10	1	3	0,031	0,177	-0.18	0,429
11	18	56	0.563	0,504	-1.12	0,131
12	32	100	1.0	0	0	***
13	32	100	1.0	0	0	***
14	32	100	1.0	0	0	***

n:	Observationsværdi
%:	Procent
X:	Den aritmetiske middelværdi
S:	Standardafvigelsen
Z:	Observationsværdiens afvigelse fra den aritmetiske middelværdi udtrykt i standardafvigelsesenheder
P:	Sandsynlighed for H_0
*:	$P < 0,05$
***:	$P < 0,001$

Som det fremgår af figur 2.3.3 er både danske og finske forsøgspersoners identifikationer statistisk signifikante (dvs. $P < 0,05$) i 12 tilfælde (markeret '*'), men de statistisk signifikante identifikationer er fordelt på forskellige stimuli: hos danske forsøgspersoner er stimulus 1-3 identificeret som S, 6-10 som T og 11-14 som D, hvor finske forsøgspersoner har identificeret stimulus 1-5 som S, 7-10 som T og 12-14 som D.

Vi kan foreløbigt konkludere følgende:

BEST COPY AVAILABLE

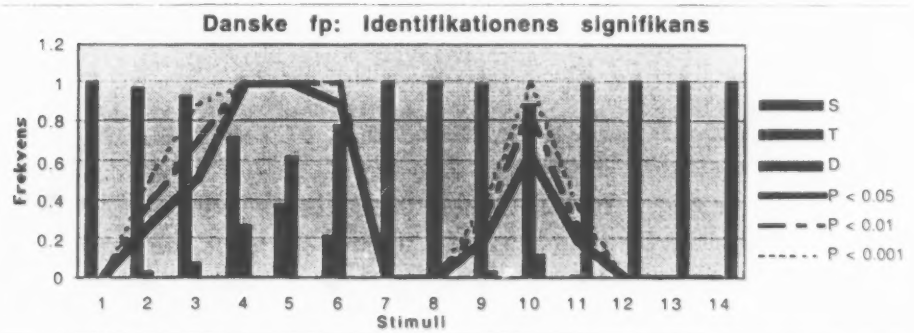
1) De finske forsøgspersoner tilskriver færre stimuli til 'D end de danske forsøgspersoner, f.eks. identificeres stimulus 11 ikke konsekvent som D af finske, men af danske forsøgspersoner.

2) De finske forsøgspersoner tilskriver færre stimuli til 'T end de danske forsøgspersoner, f.eks. identificeres stimulus 6 ikke konsekvent som T af finske, men af danske forsøgspersoner.

3) Omvendt tilskriver de finske forsøgspersoner langt flere stimuli til 'S end de danske forsøgspersoner, f.eks. identificeres stimulus 4 og 5 som S af de finske forsøgspersoner, hvorimod de danske forsøgspersoner oftere identificerer stimulus 5 som T.

Det, vi foreløbigt har kunnet sige om danske og finske forsøgspersoners identifikation af samme akustiske VOT-kontinuum, bygger på data, der er statistisk signifikante, dvs. $H_0(P < 0,05)$.

Nu kunne det være godt at vide, hvor signifikante de signifikante data er, og hvor langt de insignifikante data er fra at være signifikante. Er de insignifikante identifikationer – f.eks. de finske identifikationer af stimulus 11 eller danske identifikationer af stimulus 4 og 5 – fuldstændig tilfældige eller er de relativt tæt ved at være signifikante? Dette spørgsmål har jeg forsøgt at besvare ved at sammenligne antallet af identifikationer for hver stimulus med grænsen for hvor få identifikationer, der skal til for H_0 med hhv. $P < 0,05$, $P < 0,01$ og $P < 0,001$ (jf. figur 2.3.4).



BEST COPY AVAILABLE

Figur 2.3.4

Som vi nu kan se (i figur 2.3.4) adskilles de tre kategorier 'S, 'T og 'D ved relativ skarpe grænser; der er altså ikke bare tale om glidende overgange fra kategori til kategori over adskillige stimuli, hvilket ville have resulteret i "fladere", bølgeformede signifikanskurver. Vi kan altså med al rimelighed sige, at identifikationsadfærden hos danske og finske forsøgspersoner afspejler kategorier med skarpe grænser og ikke prototypekategorier med gradueret medlemskab (Lakoff 1987; Rosch 1981).

Hvis vi udelukkende fokuserer på de skarpt afgrænsede kategorier, ser vi, at de fordeler sig forskelligt hos de to fp-grupper:

	S	T	D
Kategorifordelingen hos danske fp (P < 0,001):	1 - 3	7 - 9	11 - 14
Kategorifordelingen hos finske fp (P < 0,001):	1 - 4	8 - 10	12 - 14

Figur 2.3.5

Set i forhold til VOT, der jo er den varierende faktor i det akustiske kontinuum, som de to grupper er blevet præsenteret for, fremgår det, at den finske D-kategori består af stimuli med VOT under 25 ms, hvorimod den danske D-kategori inkluderer stimulus 11 med VOT på 25 ms. Samme VOT-forskydning finder vi i de to T-kategorier: igen inkluderer den danske kategori stimuli med højere VOT end den tilsvarende finske kategori.

	S	T	D
Danske VOT-kategorier (P < 0,001):	125-105	65-45	25--5
Finske VOT-kategorier (P < 0,001):	125-95	55-35	15--5

Figur 2.3.6

BEST COPY AVAILABLE

Sammenligner vi de to fp-gruppers S-kategorier, opdager vi, at de følger det overordnede mønster: finske kategorier inkluderer stimuli med lavere VOT end de korresponderende danske kategorier. Foreløbig tyder alt altså på, at kategorigrænserne er sprogspecifikke, dvs. betinget af en strukturering af det akustiske VOT-kontinuum på et konkret sprog.

Ovenstående resultater svarer ganske godt til mine forventninger. Det forholder sig nemlig sådan, at den fonetiske kontrast mellem [d] og [t] er markeret på forskellig måde på hhv. finsk og dansk. Som tidligere beskrevet (side 52ff) adskiller dansk [d] og [t] sig fra hinanden udelukkende ved aspiration; hvor [d] er uaspireret, er [t] kraftigt aspireret. Heroverfor er både det finske [d] og [t] uaspireret. Derimod er kontrasten markeret ved en forskel i stemthed; på finsk er [d] en stemt lyd og [t] en ustemt lyd.

Når de finske forsøgspersoner udelukkende identificerer stimuli med ganske kort VOT som D, skyldes det netop, at varigheden af friktionsstøj før VO ikke må overskride varigheden af den eksplosionsstøj, der er en naturlig følge af artikulationsorganernes "træghed" – ofte indledes selv stemte sproglyde af en kortvarig eksplosionsfase, hvilket giver VO over 0 ms.

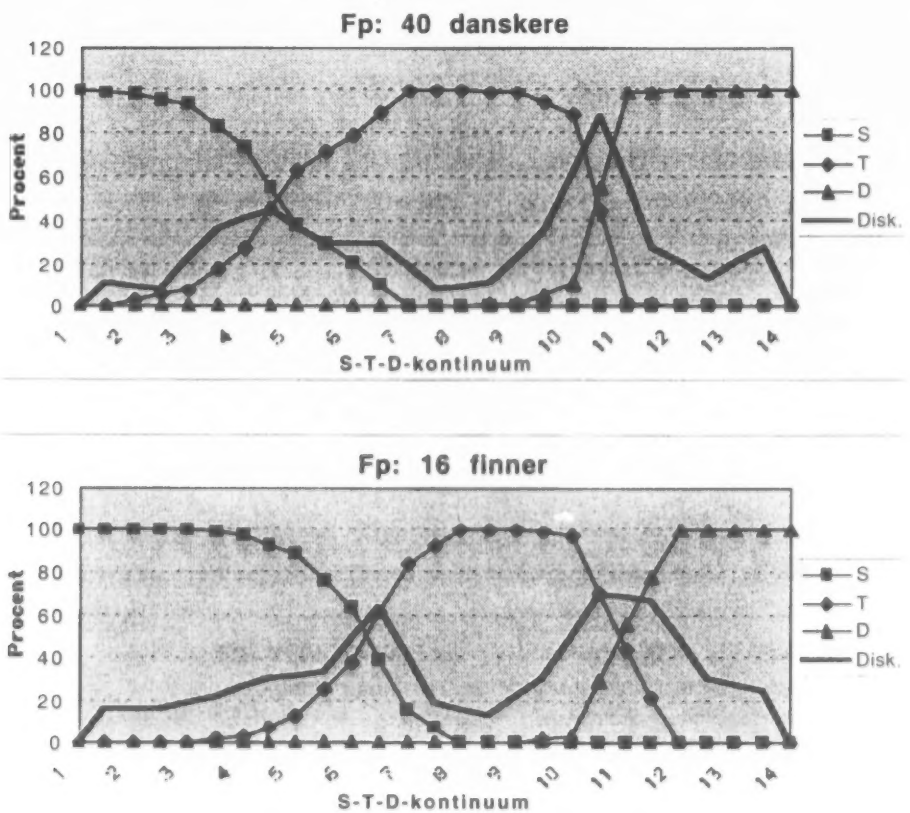
Omvendt overrasker det mig, at de finske forsøgspersoner har identificeret stimuli med en relativ lang VOT som tilhørende 'T. Det var egentlig min forventning, at det ustemte og uaspirerede [t] på finsk i højere grad korresponderede med nogle af de stimuli, som danske forsøgspersoner har identificeret som D, der på dansk ligeledes er ustemt og uaspireret. Når dette ikke er tilfældet, skyldes det sandsynligvis, at de fonetiske termer 'stemthed' og 'aspiration' dækker over en relativ stor variation i konkrete manifestationer.

Diskrimination af stimuli fra et VOT-kontinuum

Nu er det væsentligste kriterium for kategorisk perception ikke identifikationen alene, men dens relation til diskriminationen af samme stimuli. Data fra en lang række forsøg med KP viser dog, at der sjældent er tale om ideel kategorisk perception, dvs. en diskriminationsadfærd, der udelukkende er betinget af kategoriseringen. Oftest er diskriminationen af stimuli fra samme kategori simpelthen bedre end Haskins-modellens forudsigelser.

Ovenstående analyse af resultater fra identifikationsserierne afslører et markant skifte i identifikationen som funktion af VOT. På baggrund af disse resultater forventer jeg at finde en tilsvarende god diskrimination af stimuli fra forskellige kategorier, dvs. med en maksimal diskrimination ved kategori-grænserne, samt en markant ringe diskrimination af stimuli fra samme kategori.

Figur 2.3.7 nedenfor bygger på resultater fra de 4 diskriminationsserier, hvor forsøgspersonerne skulle vurdere om to præsenterede (nabo-)stimuli var ens eller forskellige (se beskrivelsen af diskriminationsserier og procedure side 67ff).



Figur 2.3.7

Gruppen af danske fp bestod af 40 personer, der hver 40 gange skulle afgive en 'ens/forskellige'-vurdering af de præsenterede stimuli, hvilket betyder, at

fordelingen i figur 2.3.7 er baseret på 1.600 præsentationer og vurderinger (40 · 40).

Den finske fp-gruppe bestod af 16 personer, der ligeledes skulle afgive hver 40 'ens/forskellige'-vurderinger, hvilket gav i alt 640 præsentationer og vurderinger (16 · 40).

Som det med al tydelighed fremgår af figur 2.3.7 er der en nær korrelation mellem identifikationen og diskriminationen (begge som funktion af VOT). For det første er diskriminationen af stimuli fra samme kategori meget lav. Den højeste diskriminationsprocent mellem stimuli fra samme kategori er hos de danske forsøgspersoner 36% (mellem stimulus 3 og 4), dvs. et ringere udfald end hvis forsøgspersonerne havde slået plat og krone[2]. Derimod er den tilsvarende højeste diskriminationsprocent hos de finske forsøgspersoner lidt højere, nemlig 66% (mellem stimulus 11 og 12).

For det andet kan vi konstatere en relativ god diskrimination af stimuli fra forskellige kategorier – dvs. ved kategorigrænser og -overgange. Hos de danske forsøgspersoner kan vi notere en diskriminationsprocent helt oppe på 88% (mellem stimulus 10 og 11), hvilket betyder, at de i knapt 9 ud af 10 præsentationer har kunnet høre forskel på stimuli varierende med 10 ms VOT. De to diskriminationstoppe stemmer meget godt overens med de tidligere anslåede kategorigrænser ved hhv. stimulus 4-5 (VOT i ms: 95/85) og stimulus 10-11 (VOT i ms: 35/25). Begge diskriminationstoppe er nærmest prototypiske eksempler på fænomenet 'kategorisk effekt'.

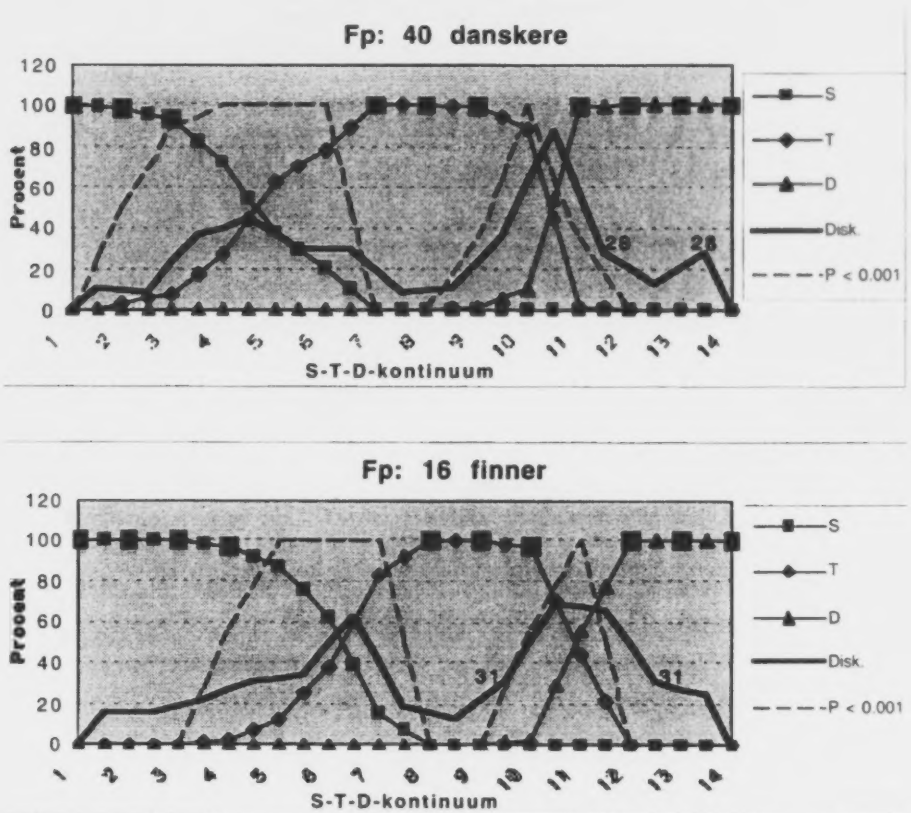
De finske forsøgspersoner har ligeledes haft lettest ved at skelne mellem stimulus 10 og 11 – dog kun i 69% af præsentationerne. Den anden kategorigrænse, der ligeledes er manifesteret ved en diskriminationstop, findes som forventet mellem stimulus 6 og 7 (VOT i ms: 75/65). Begge diskriminationstoppe er igen glimrende eksempler på den kategoriske effekt.

Der er altså ingen tvivl om, at vi på grundlag af disse data kan konkludere, at såvel de danske som de finske resultater afspejler kategorisk perception.

[2] Og her har jeg endda medregnet stimuli, hvis identifikation er signifikant ved $P < 0,05$ og ikke kun de skarptafgrænsede kategorier med $P < 0,001$.

Statistiske analyser

Hvis vi herefter udelukkende fokuserer på diskriminationen af stimuli, der er signifikante med $P < 0{,}001$, træder et endnu klarere mønster frem. I figur 2.3.8 har jeg sammenholdt de to fp-gruppers identifikationer, identifikationernes signifikansgrænse for $H_0(P < 0{,}001)$ samt diskrimination. Den stiplede signifikanskurve for $P < 0{,}001$ afgrænser de 10 mest signifikante stimuli, som jeg for overskuelighedens skyld har markeret med store firkantede koordinater.



Figur 2.3.8

De 10 stimuli, der fordeler sig over tre kategorier, er ikke alene meget signifikante, men også de sværeste at diskriminere. Hos de danske forsøgspersoner har de 10 mest signifikante stimuli de laveste diskriminationsprocenter med 28% som det maksimale, jf. de markerede punkter på

BEST COPY AVAILABLE

diskriminationskurven. Det er især værd at lægge mærke til den påfaldende stigning i diskrimination af stimulus 13 og 14, hvilket skyldes, at dette interval overskrider stemthedsgrænsen ved 0 ms VOT.

Hos den finske fp-gruppe udgør de 10 mest signifikante stimuli ligeledes rækken af stimuli, der er sværest at diskriminere – dog med en enkelt undtagelse: diskriminationen af stimulus 7 og 8 er på 19% – altså lavere end mellem f.eks. 3 og 4, 9 og 10, 12, 13 og 14 – hvilket sandsynligvis skyldes, at stimulus 8 er signifikant med $P < 0{,}001$ og stimulus 7 med $P < 0{,}01$. Som det fremgår af de markerede punkter på diskriminationskurven er den maksimale diskriminationsprocent for de mest signifikante stimuli på 31% .

Som ovenstående analyser illustrerer, afspejler både danske og finske resultater kategorisk perception. Den bedste indikator for KP finder vi i den kategoriske effekt, dvs. diskriminationstoppene mellem kategorier.

Perspektivering af resultater

Spørgsmålet om hvorvidt den kategoriske effekt skyldes en psykoakustisk tærskel eller sprogspecifik fonetisk kategorisering kan ikke besvares éntydigt. Det er dog helt sikkert, at forsøgspersonens modersmål har en meget stor indflydelse på såvel identifikationen som diskriminationen af stimuli fra et VOT-kontinuum. Det fremgår klart af resultaterne, at de to fp-grupper strukturerer det akustiske kontinuum forskelligt – helt i overensstemmelse med de fonetiske forskelle på de to sprog.

Omvendt kan fonetisk betinget kategorisering ikke forklare alle fænomener: f.eks. er den øgede diskrimination af stimulus 13 og 14 hos de danske forsøgspersoner ikke betinget af en fonetisk stemthedskontrast på det pågældende sprog. På trods af at stemthedsgrænsen (0 ms VOT) ikke indgår i den danske strukturering af udtrykssubstansen, er de danske forsøgspersoner i stand til at skelne mellem stimulus 13 og 14 med hhv. +5 ms VOT og -5 ms VOT i 28% af tilfældene. Selvom sprogperceptionen i høj grad er betinget af sprogspecifikke kategorier, er der altså ikke belæg for at gå så langt som til at sige, at:

"...discrimination is limited by identification: subjects can only discriminate between stimuli that they identify differently" (Studdert-Kennedy, Liberman, Harris & Cooper 1970: 234).

98

En anden problemstilling, jeg har forsøgt at belyse i dette forsøg med KP, er spørgsmålet om, hvorvidt der eksisterer en nær sammenhæng mellem artikulation og perception, som postuleret af Haskins-gruppen o.a. Ifølge Haskins-modellen udspringer sprogproduktion og sprogperception af samme sæt artikulatoriske kommandoer. Hvis denne antagelse er korrekt, medfører det, at sprogbrugere ikke har adgang til information ud over de fonetiske kategorier, dvs. at fonetiske kategorier, der adskiller sig ved diskrete artikulatoriske egenskaber, perciperes som absolutte, diskrete størrelser, hvorimod perceptionen er kontinuerlig, når kontinuerlig artikulatorisk variation er mulig.

Artikulatorisk kontinuitet er efter min opfattelse ikke altid et spørgsmål om 'enten-eller', men 'mere' eller 'mindre'. F.eks. kan artikulationen kaldes kontinuerlig, hvis en sprogbruger er i stand til at producere vokallyde, der udgør forskellige kvaliteter mellem f.eks. et [æ] og et [a]. Omvendt er den artikulatoriske overgang mindre kontinuerlig ved produktionen af sproglyde, der udgør overgange fra et [t] til et [d] – her har danske sprogbrugere langt sværere ved at producere tilsvarende små forskelle – men slet ikke lige så diskontinuerlig som overgangen fra f.eks. et [t] til et [p][3]. Set i dette perspektiv udgør de tre fonetiske kategorier, som indgår i dette forsøg, hhv. en artikulatorisk diskontinuerlig overgang fra [s] til [t], men også en overgang fra [t] til [d], der hverken er absolut diskontinuerlig eller kontinuerlig.

Med afsæt i synteseteoriernes antagelse om at KP afspejler menneskers artikulatoriske potentiale, kan vi redefinere motorhypotesen, så den tager hensyn til den beskrevne graduering af 'artikulatorisk kontinuitet': jo mindre artikulatorisk kontinuerlige sproglydene er, desto større kategorisk effekt vil de udvise – og større kontinuitet vil resultere i mindre kategorisk effekt. Fordelen ved denne redefinition er, at motorhypotesen nu kan opstilles, så den er empirisk falsificerbar: Hvis artikulatorisk kontinuitet er den eneste faktor, der betinger KP, vil diskriminationen af akustisk diskontinuerlige stimuli være bedre end af akustisk kontinuerlige stimuli. Ser vi nærmere på de danske forsøgspersoners evne til diskrimination, jf. figur 2.3.8, viser det sig, at den maksimale diskriminationsprocent mellem de artikulatorisk dis-

[3] Jeg har altså valgt at fortolke begrebet 'artikulatorisk kontinuitet' som evnen til at producere sproglyde, der udgør overgange mellem to fonetiske kategorier. Det er dog også muligt at hævde, at 'artikulatorisk kontinuitet' består i muligheden for at producere en sproglyd, der i sig selv udgør en artikulatorisk overgang mellem fonetiske kategorier. Der er nok ingen tvivl om, at Haskins-gruppen henviser til sidstnævnte definition.

kontinuerlige kategorier 'S og 'T er 45%, altså væsentlig lavere end den maksimale diskrimination på 88% mellem kategorierne 'T og 'D, der er mere artikulatorisk kontinuerlige.

Der er altså ikke hjemmel i mine empiriske data for at hævde en én-til-én relation mellem artikulation og perception. Selvom data fra de to fp-grupper klart afspejler kategorisk perception, ville en perception udelukkende baseret på artikulatoriske kommandoer have givet et helt andet billede. For det første ville identifikationsfunktionen for hver stimulus have været enten 100% eller 0%. For det andet – og her finder vi den største afvigelse – ville diskriminationen af stimuli fra to artikulatorisk diskontinuerlige kategorier have været 100% og ikke 45%. Endelig skulle forsøgspersonerne ikke have adgang til information ud over de fonetiske kategorier, hvilket betyder, at diskriminationen af stimuli fra samme kategori ville være på 0% og ikke oppe på 28%, som vi kan notere hos de danske forsøgspersoner. Forsøgspersonerne må altså også have mulighed for at benytte anden information end de postulerede artikulatoriske kommandoer.

2.4 Resultater fra forsøg med kategorisk perception: et I-E-Æ-A-kontinuum

I det følgende vil jeg præsentere resultaterne fra den del af forsøget med KP, der involverer et akustisk vokalkontinuum. De præsenterede stimuli er syntetisk produceret med varierende F1 og F2 (se beskrivelsen af stimuli side 59ff). Resultaterne bygger på identifikations- og diskriminationsadfærd hos de tidligere omtalte fp-grupper, jf. side 71ff.

De er min hensigt at præsentere resultaterne i samme rækkefølge, som jeg præsenterede resultaterne fra forsøg med et VOT-kontinuum. De første resultater, jeg vil gennemgå, er de to fp-gruppers identifikation af stimuli. Disse resultater vil jeg analysere med samme statistiske metoder som blev anvendt i analysen af konsonantidentifikationer i foregående kapitel. Derefter vil jeg præsentere resultaterne fra diskriminationsserierne og analysere deres forhold til identifikationen for at finde ud af, om de samlet opfylder kriterierne for KP.

Endelig vil jeg perspektivere resultaterne med henblik på de opstillede hypoteser i kapitel 2.2: Forsøg med kategorisk perception, nemlig motor-hypotesen, hypotesen om sprogspecifikke kategorigrænser og hypotesen om stimulusasymmetri.

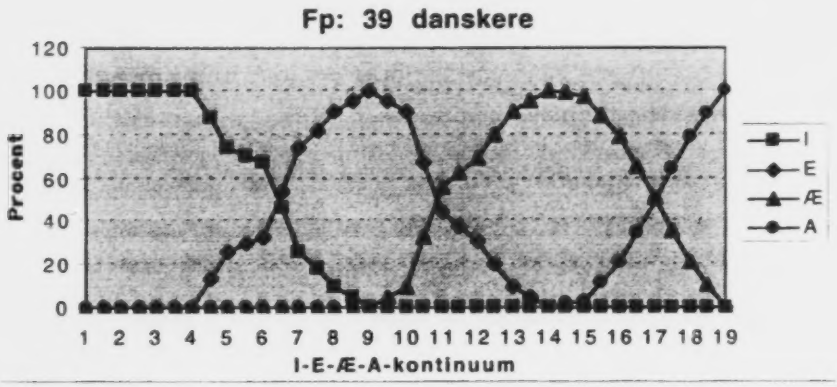
Identifikation af stimuli fra et formantkontinuum

De første resultater, jeg vil præsentere, er fra de 4 identifikationsserier, hvor forsøgspersonerne skulle identificere de 19 præsenterede stimuli som enten 'A', 'E', 'I' eller 'Æ'.

Den danske fp-gruppe bestod i dette forsøg "kun" af 39 personer (jf. fp-oversigten side 73ff). Forsøgspersonerne blev i dette forsøg præsenteret for en identifikationsserie bestående af 19 præsentationer, hvilket betyder, at procentfordelingen i figur 2.4.1 er baseret på 741 præsentationer og identifikationer (39 · 19).

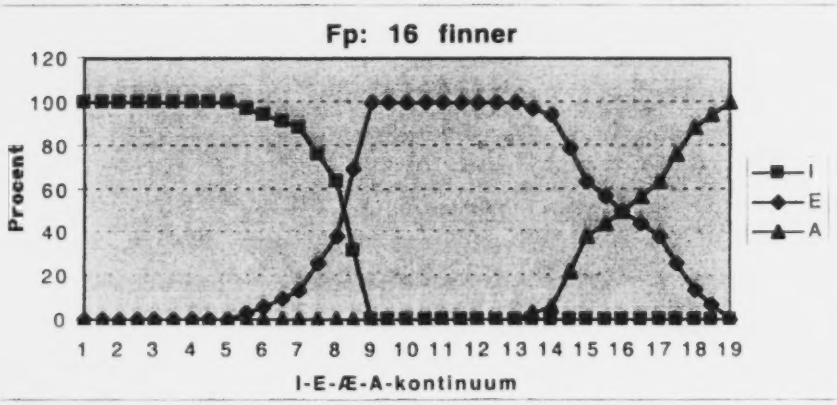
Som fordelingskurven tydeligt illustrerer, har forsøgspersonerne kunnet identificere stimuli som tilhørende én af de fire kategorier, 'I, 'E, 'Æ eller 'A. Den leptokurtiske fordeling, dvs. en mere spids fordelingstop end ved normalfordelingen, antyder, at forsøgspersonerne kun har fundet et lille antal

stimuli egnet som eksempler på naturlige vokalmanifestationer. Omvendt har disse stimuli udgjort "gode" eksempler, for de er blevet tilskrevet samme kategori hver gang.



Figur 2.4.1

I histogrammet over de finske forsøgspersoners identifikation af samme akustiske kontinuum, figur 2.4.2, finder vi samme høje identifikationsprocent, men en anden fordelingskurtose end i det danske. Den platykurtiske fordelingsform i det finske diagram antyder, enten at de finske kategorier dækker større frekvensområder, eller at de finske forsøgspersoner er mere fonetisk fleksible end de danske.



Figur 2.4.2

102

Gruppen af finske forsøgspersoner bestod af de samme 16 personer, som indgik i forsøget med konsonantstimuli. De 16 finske forsøgspersoner blev ligesom de danske præsenteret for de 19 syntetiske stimuli, som de skulle identificere med afsæt i deres eget modersmål. Procentfordelingen i figur 2.4.2 er altså baseret på 304 præsentationer og identifikationer (16 · 19).

Som det fremgår af fordelingskurven har de finske forsøgspersoner kun identificeret de præsenterede stimuli som I, E og A – ikke Æ. Dette skyldes, at der kun eksisterer tre urundede fortungevokaler på finsk, et [i] og et [e], der begge artikuleres med lidt større åbningsgrad end de tilsvarende danske. Desuden har finsk – som i dansk – to a-varianter, en urundet fortungevokal, der er en anelse mere åbne end det danske [æ] (som f.eks. i 'abe') og en bagtungevariant.

Som følge af denne forskydning mellem de danske og de finske vokallyde opstod der (uventet) variation i de finske valg af respons. Nogle af de finske forsøgspersoner valgte en skrift-til-lyd strategi, så det danske vokalbogstav 'æ' repræsenterede den finske fortungevokal og 'a' bagtungevokalen. Hos disse forsøgspersoner blev det syntetiske vokalkontinuum tilskrevet de tre kategorier 'I, 'E og 'Æ, hvorimod 'A blev udeladt. Andre af de finske forsøgspersoner valgte derimod en lyd-til-lyd strategi med 'a' som repræsentation for den finske fortungevokal. Hos disse forsøgspersoner blev ingen af de præsenterede stimuli altså tilskrevet kategorien 'Æ. For at undgå at inddrage fire finske kategorier af urundede fortungevokaler, hvor der faktisk kun eksisterer tre, har jeg valgt at se bort fra forskellen mellem respons med 'æ' og respons med 'a'. De to fiktive kategorier 'Æ og 'A har jeg altså slået sammen til én kategori, der repræsenteres ved et 'A'.

Ser vi på, hvordan de to fp-grupper har tilskrevet kategorier til det akustiske kontinuum, træder der nogle klare forskelle frem. Hvis vi udelukkende fokuserer på de observationsværdier, der overstiger 50% percentilen[1] (jf. figur 2.4.1 & 2), kan vi se, at de danske forsøgspersoner har tilskrevet stimulus 1-6 til kategorien 'I, 7-10 til 'E, 11-17 til 'Æ og 18-19 til 'A. Tilsvarende har de finske forsøgspersoner tilskrevet stimulus 1-8 til 'I, 9-15 til 'E og 17-19 til 'A.

[1] 50% percentilen er, som omtalt i note 1, side 77, ikke udtryk for medianen (Petersen 1995). Vi fokuserer her udelukkende på de observationsværdier, der overstiger 50% percentilen, hvilket ekskluderer de finske identifikationer af stimulus 16, der i præcis 50% af tilfældene er blevet identificeret som E og i de andre 50% som A.

De danske forsøgspersoner har altså kategorier, der generelt inkluderer færre stimuli end de finske forsøgspersoner. En af årsagerne til dette er naturligvis, at det akustiske kontinuum strækker sig over fire danske kategorier, men kun tre finske. En anden mulig årsag er, at de finske forsøgspersoner accepterer langt større variation, først og fremmest i E-kategorien. De meget skarpe kategorigrænser i det finske histogram – først og fremmest mellem 'I og 'E – tyder på en højere grad af kategorisk perception end det danske diagram indikerer.

Statistiske analyser

Lad os nu se på, om kategoritilskrivningen kan være opstået ved en tilfældighed: alt andet lige er der 25% sandsynlighed for, at de danske forsøgspersoner har identificeret en vilkårlig stimulus som tilhørende én af de fire kategorier. Tilsvarende er der 33% sandsynlighed ved de finske identifikationer, der jo indbefatter tre kategorier.

På samme måde som i analyserne af konsonantidentifikationer har jeg opstillet en statistisk 'nul-hypotese', nemlig at en given stimulus ikke identificeres som tilhørende en given kategori. Sandsynlighed for H_0 for hver enkelt stimulus fremgår af P-værdierne i figur 2.4.3. Sandsynligheden er beregnet ud fra forholdet mellem observationsværdien (n), populationens aritmetiske middelværdi (X) og standardafvigelse (S) set i forhold til en normalfordeling (den gaussiske fordeling). Sandsynligheden for H_0 er ved nogle stimuli så lav, at den ikke indgår i min normalfordelingstabel (med P som funktion af Z) (Petersen 1995). Disse sandsynligheder ($P < 0{,}001$) er i figur 2.4.3 markeret med tre stjerner (***). $H_0(P < 0{,}05)$, dvs. en statistisk signifikant fordeling, er derimod markeret med én stjerne (*).

Population: danske forsøgspersoners identifikationer
Kategori: I

Stimulus	n	%	X	S	Z	P
1	39	100	1,0	0	0	***
2	39	100	1,0	0	0	***
3	39	100	1,0	0	0	***
4	39	100	1,0	0	0	***
5	29	74	0,744	0,442	-1,683	0,046*
6	26	67	0,667	0,478	-1,395	0,081
7	10	26	0,256	0,442	-0,579	0,281

8	4	10	0.103	0,307	0.336	0.367
9	0	0	0	0	0	1,0
10	0	0	0	0	0	1,0
11	0	0	0	0	0	1,0
12	0	0	0	0	0	1,0
13	0	0	0	0	0	1,0
14	0	0	0	0	0	1,0
15	0	0	0	0	0	1,0
16	0	0	0	0	0	1,0
17	0	0	0	0	0	1,0
18	0	0	0	0	0	1,0
19	0	0	0	0	0	1,0

Kategori: E

Stimulus	n	%	X	S	Z	P
1	0	0	0	0	0	1,0
2	0	0	0	0	0	1,0
3	0	0	0	0	0	1,0
4	0	0	0	0	0	1,0
5	10	26	0.256	0,442	-0,579	0,281
6	13	33	0.333	0,478	-0,697	0,242
7	29	74	0,744	0,442	-1.683	0,046*
8	35	90	0,897	0,307	-2.922	0,002*
9	39	100	1,0	0	0	***
10	35	90	0.897	0,307	-2.922	0,002*
11	17	44	0.436	0,502	-0.869	0,197
12	12	31	0.308	0,468	-0.658	0,255
13	4	10	0,103	0,307	-0.336	0,367
14	0	0	0	0	0	1,0
15	0	0	0	0	0	1,0
16	0	0	0	0	0	1,0
17	0	0	0	0	0	1,0
18	0	0	0	0	0	1,0
19	0	0	0	0	0	1,0

BEST COPY AVAILABLE

Kategori: Æ

Stimulus	n	%	X	S	Z	P
1	0	0	0	0	0	1.0
2	0	0	0	0	0	1,0
3	0	0	0	0	0	1,0
4	0	0	0	0	0	1,0
5	0	0	0	0	0	1,0
6	0	0	0	0	0	1,0
7	0	0	0	0	0	1,0
8	0	0	0	0	0	1,0
9	0	0	0	0	0	1,0
10	4	10	0.103	0,307	-0.336	0,367
11	22	56	0.564	0,502	-1.124	0,131
12	27	69	0.692	0,468	-1.479	0,069
13	35	90	0.897	0,307	-2.922	0,002*
14	39	100	1.0	0	0	***
15	38	97	0.974	0,160	-6.088	***
16	31	79	0.795	0,409	-1.944	0,026*
17	20	51	0.513	0,506	-1.014	0,156
18	8	21	0.205	0,409	-0.501	0,308
19	0	0	0	0	0	1.0

Kategori: Å

Stimulus	n	%	X	S	Z	P
1	0	0	0	0	0	1,0
2	0	0	0	0	0	1,0
3	0	0	0	0	0	1,0
4	0	0	0	0	0	1,0
5	0	0	0	0	0	1,0
6	0	0	0	0	0	1,0
7	0	0	0	0	0	1,0
8	0	0	0	0	0	1,0
9	0	0	0	0	0	1,0
10	0	0	0	0	0	1,0
11	0	0	0	0	0	1,0
12	0	0	0	0	0	1,0

BEST COPY AVAILABLE

13	0	0	0	0	0	1,0
14	0	0	0	0	0	1,0
15	1	3	0.026	0,160	-0.163	0,436
16	8	21	0.205	0,409	-0.501	0,308
17	19	49	0.487	0,506	-0.962	0,168
18	31	79	0.795	0,409	-1.944	0,026*
19	39	100	1.0	0	0	***

Population: finske forsøgspersoners identifikationer

Kategori: I

Stimulus	n	%	X	S	Z	P
1	16	100	1.0	0	0	***
2	16	100	1.0	0	0	***
3	16	100	1.0	0	0	***
4	16	100	1.0	0	0	***
5	16	100	1.0	0	0	***
6	15	94	0.938	0,250	-3.75	***
7	14	88	0.875	0,342	-2.56	0,005*
8	10	63	0.625	0,500	-1.25	0,106
9	0	0	0	0	0	1,0
10	0	0	0	0	0	1,0
11	0	0	0	0	0	1,0
12	0	0	0	0	0	1,0
13	0	0	0	0	0	1,0
14	0	0	0	0	0	1,0
15	0	0	0	0	0	1,0
16	0	0	0	0	0	1,0
17	0	0	0	0	0	1,0
18	0	0	0	0	0	1,0
19	0	0	0	0	0	1,0

Kategori: E

Stimulus	n	%	X	S	Z	P
1	0	0	0	0	0	1,0
2	0	0	0	0	0	1,0
3	0	0	0	0	0	1,0
4	0	0	0	0	0	1,0

BEST COPY AVAILABLE

5	0	0	0	0	0	1,0
6	1	6	0,063	0,250	-0,25	0,401
7	2	13	0,125	0,342	-0,37	0,356
8	6	38	0,375	0,500	-0,75	0,227
9	16	100	1,0	0	0	***
10	16	100	1,0	0	0	***
11	16	100	1,0	0	0	***
12	16	100	1,0	0	0	***
13	16	100	1,0	0	0	***
14	15	94	0,938	0,250	-3,75	***
15	10	63	0,625	0,500	-1,25	0,106
16	8	50	0,5	0,516	-0,97	0,166
17	6	38	0,375	0,500	-0,75	0,227
18	2	13	0,125	0,342	-0,37	0,356
19	0	0	0	0	0	1,0

Kategori: A

Stimulus	n	%	X	S	Z	P
1	0	0	0	0	0	1,0
2	0	0	0	0	0	1,0
3	0	0	0	0	0	1,0
4	0	0	0	0	0	1,0
5	0	0	0	0	0	1,0
6	0	0	0	0	0	1,0
7	0	0	0	0	0	1,0
8	0	0	0	0	0	1,0
9	0	0	0	0	0	1,0
10	0	0	0	0	0	1,0
11	0	0	0	0	0	1,0
12	0	0	0	0	0	1,0
13	0	0	0	0	0	1,0
14	1	6	0,063	0,250	-0,25	0,401
15	6	38	0,375	0,500	-0,75	0,227
16	8	50	0,500	0,516	-0,97	0,166
17	10	63	0,625	0,500	-1,25	0,106
18	14	88	0,875	0,342	-2,56	0,005*

BEST COPY AVAILABLE

19	16	100	1.0	0	0	***

n:	Observationsværdi
%:	Procent
X:	Den aritmetiske middelværdi
S:	Standardafvigelsen
Z:	Observationsværdiens afvigelse fra den aritmetiske middelværdi udtrykt i standardafvigelsesenheder
P:	Sandsynlighed for H_0
*:	$P < 0.05$
***:	$P < 0.001$

Figur 2.4.3

Som det fremgår af figur 2.4.3 er både danske og finske forsøgspersoners identifikationer statistisk signifikante (dvs. $P < 0,05$) i 15 tilfælde (markeret *), men de statistisk signifikante identifikationer er fordelt på forskellige stimuli: hos danske forsøgspersoner er stimulus 1-5 identificeret som I, 7-10 som E, 13-16 som Æ og 18-19 som A, hvor finske forsøgspersoner har identificeret stimulus 1-7 som I, 9-14 som E og 18-19 som A.

Vi kan foreløbig konkludere følgende:
1) De danske forsøgspersoner tilskriver færre stimuli til 'I end de finske forsøgspersoner, f.eks. identificeres stimulus 6 ikke konsekvent som I af danske, men af finske forsøgspersoner. Endvidere identificerer danske forsøgspersoner stimulus 7 som E, hvorimod finske forsøgspersoner identificerer samme stimulus som I. Stimulus 7 indgår altså i to forskellige kategorier på de to sprog.

2) De danske forsøgspersoner tilskriver ligeledes færre stimuli til 'E end de finske forsøgspersoner, f.eks. identificeres stimulus 11 og 12 ikke som E af danske, men af finske forsøgspersoner. Endvidere identificerer danske forsøgspersoner stimulus 13 og 14 som Æ, hvorimod finske forsøgspersoner identificerer de samme stimuli som E. Stimulus 13 og 14 indgår altså også i to forskellige kategorier på de to sprog.

3) Hvor de danske forsøgspersoner tilskriver stimulus 13-16 til 'Æ, indgår denne kategori ikke i de finske identifikationer. Det lader til, at denne

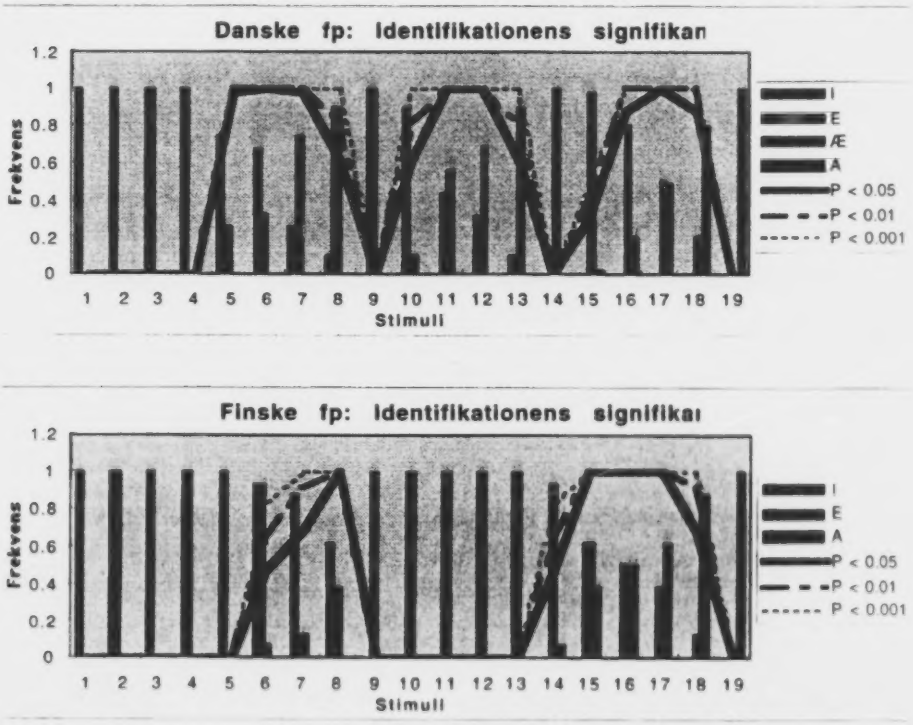
BEST COPY AVAILABLE

kategori er blevet assimileret af 'E, da stimulus 13, 14 og 15 oftere bliver identificeret med E end A og stimulus 16 falder lige på 50% percentilen.

4) Endelig er der sammenfald mellem de to sproggruppers A-kategori; begge fp-grupper identificerer stimulus 18 og 19 som A.

Det, vi foreløbigt har kunnet sige om danske og finske forsøgspersoners identifikation af samme akustiske F1/F2-kontinuum, bygger på data, der er statistisk signifikante, dvs. $H_0(P < 0,05)$.

Nu kunne det være godt at vide, hvor signifikante de signifikante data er, og hvor langt de insignifikante data er fra at være signifikante. Dette spørgsmål har jeg forsøgt at besvare ved at sammenligne antallet af identifikationer for hver stimulus med grænsen for hvor få identifikationer, der skal til for H_0 med hhv. $P < 0,05$, $P < 0,01$ og $P < 0,001$ (jf. figur 2.4.4).



Figur 2.4.4

Som det fremgår af figur 2.4.4 adskilles de fire danske ('I, 'E, 'Æ og 'A) og tre finske kategorier ('I, 'E og 'A) ved relativt skarpe grænser, dvs., at for langt de fleste stimuli gælder det, at de enten er inkluderet i eller ekskluderet af en given kategori. Der er altså ingen kraftige tendenser til glidende overgange fra kategori til kategori, hvilket ville have resulteret i fladere signifikanskurver centreret omkring 50% percentilen. Hvis vi udelukkende ser på de stimuli, hvis identifikationer er signifikante med $P < 0{,}001$, dvs. de finske identifikationer af stimulus 1-6, 9-14 og 19, og de danske identifikationer af stimulus 1-4, 9, 14-15 og 19, træder dettte mønster blot tydeligere frem. I modsætning til den forventede tendens til prototype-kategorier med gradueret medlemskab, afspejler identifikationsadfærden hos danske og finske forsøgspersoner altså relativt skarpe kategorigrænser.

Generelt dækker de finske kategorier over et større antal stimuli end de danske. Fokuserer vi på antallet af stimuli, der omfattes af I- og E-kategorien med $P < 0{,}001$, består de finske kategorier af hhv. 2 og 5 flere stimuli end de danske. Derimod lader der til at være et sammenfald mellem den danske og den finske A-kategori, som kun omfatter stimulus 19. Årsagen til kategoriernes forskellige omfang på de to sprog er naturligvis, at det samme akustiske kontinuum strækker sig over fire danske, men kun tre finske kategorier.

	I	E	Æ	A
Kategorifordelingen hos danske fp ($P < 0{,}001$):	1-4	9	14-15	19
Kategorifordelingen hos finske fp ($P < 0{,}001$):	1-6	9-14	0	19

Figur 2.4.5

Set i forhold til formantfrekvenserne, der jo er den varierende faktor i det akustiske kontinuum, fremgår det, at den finske I-kategori er karakteriseret ved andre formantfrekvenser end den tilsvarende danske kategori. Både F1 og F2 i det finske 'I dækker et større frekvensområde end de tilsvarende danske formantfrekvenser: på den ene side omfatter F1 højere frekvenser og på den anden side dækker F2 lavere frekvensområder. En rimelig tolkning af disse akustiske oplysninger må være, at finske forsøgspersoner identificerer vokallyde, der korresponderer med en mere åben kæbestilling og tilbage-trukket tungemasse end i det danske [i], som manifestationer af 'I.

	I	E	Æ	A
Danske fp (P < 0,001):	F1:200-250	F1:333	F1:417-433	F1:500
	F2:2667-2800	F2:2444	F2:2178-2222	F2:2000
Finske fp (P < 0.001):	F1:200-283	F1:333-417	F1:	F1:500
	F2:2578-2800	F2:2222-2444	F2:	F2:2000

Figur 2.4.6

Hvis vi sammenligner de to respondentgruppers E-kategorier træder nogle endnu større forskelle frem. Det danske 'E dækker her kun en lille del af det frekvensområde, der udgør det finske 'E, nemlig den laveste F1 og den højeste F2. Med andre ord har de danske forsøgspersoner kun (konsekvent) identificeret et [i], når de præsenterede vokallyde korresponderede med den mest lukkede og fremskudte manifestation af det finske 'E.

Omvendt dækker den danske Æ-kategori en stor del af de stimuli, som de finske forsøgspersoner identificerede som 'E'. Meget tyder altså på, at de formantfrekvenser, der udgør det finske 'E, korresponderer med de to danske kategorier, 'E og 'Æ.

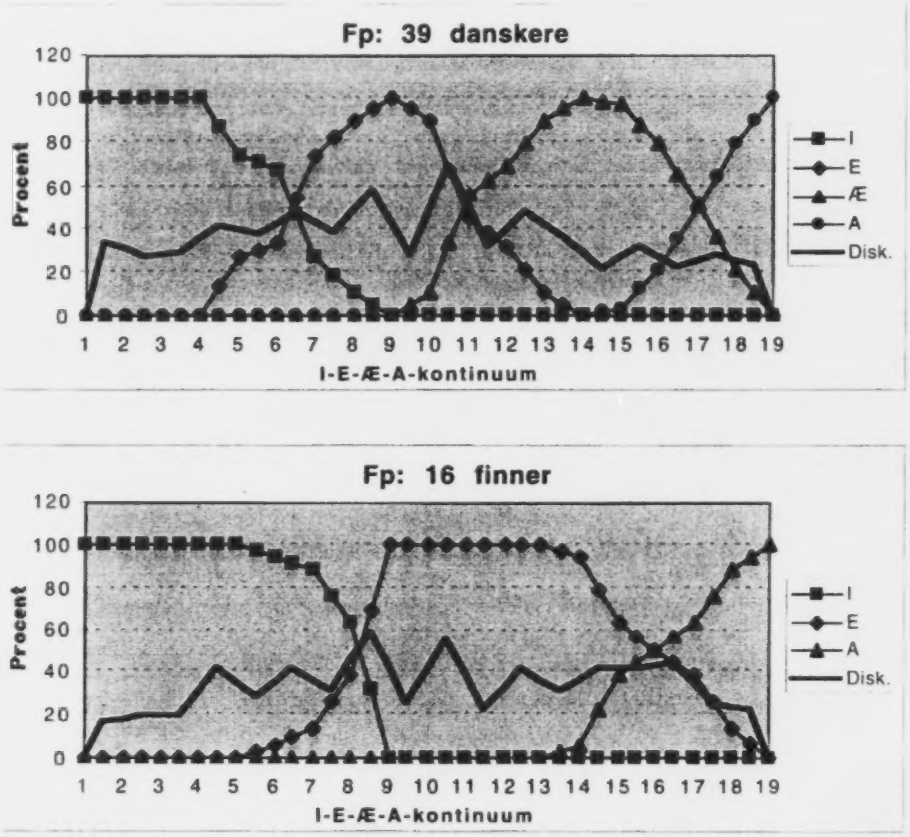
Diskrimination af stimuli fra et formantkontinuum

Ovenstående analyse af resultater fra identifikationsserierne afslører et markant skifte i identifikationen som funktion af varierende formant-frekvenser. Hvis vi udelukkende fokuserer på identifikationsfunktionen, fremstår både danske og finske forsøgspersoners vokalperception som relativt kategorisk. Identifikationen er dog ikke det eneste – eller væsentligste – kriterium for KP. KP er snarere kendetegnet ved, at forsøgspersonerne ikke kan skelne mellem flere stimuli end de kan identificere (Studdert-Kennedy et al. 1970). Med andre ord er den kategoriske effekt en nødvendig – men ikke tilstrækkelig – betingelse for, at vi overhovedet kan tale om kategorisk perception af sproglyde, dvs., at vi kan registrere en maksimal diskrimination ved kategorigrænserne, samt en markant ringe diskrimination af stimuli fra samme kategori.

Figur 2.4.7 nedenfor bygger på resultater fra de 4 diskriminationsserier, hvor forsøgspersonerne skulle vurdere om to præsenterede (nabo-)stimuli var ens eller forskellige (se beskrivelsen af diskriminationsserier og procedure side 68ff).

Gruppen af danske forsøgspersoner bestod ligesom i identifikationsøvelsen af 39 personer, der hver 55 gange skulle afgive en 'ens/forskellige'-vurdering af de præsenterede stimuli, hvilket betyder, at fordelingen i figur 2.4.7 er baseret på 2.145 præsentationer og vurderinger (39 · 55).

Den finske fp-gruppe bestod i alle forsøg af 16 personer, der ligeledes skulle afgive hver 55 'ens/forskellige'-vurderinger, hvilket gav i alt 880 præsentationer og vurderinger (16 · 55).



Figur 2.4.7

Som det fremgår af figur 2.4.7 er der kun i bedste fald tale om en ganske ringe korrelation mellem identifikationen og diskriminationen.

For det første er diskriminationen af stimuli fra samme kategori alt for god. Den højeste diskriminationsprocent mellem stimuli fra samme kategori udgør

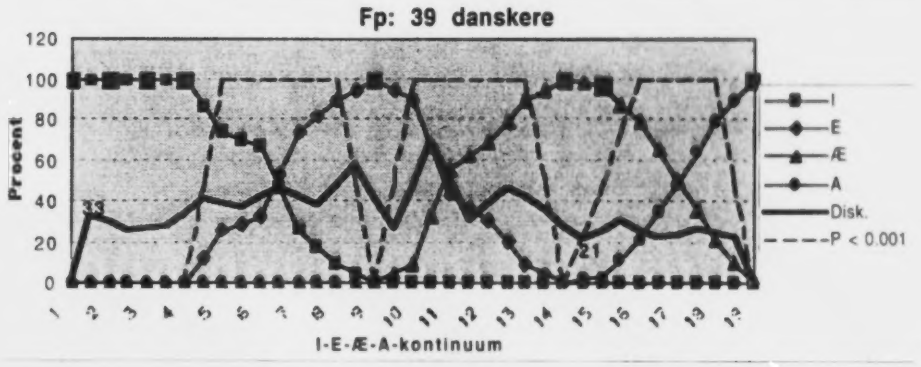
58% (mellem stimulus 8 og 9) hos de danske forsøgspersoner og 56% (mellem stimulus 10 og 11) hos de finske forsøgspersoner – begge over tilfældighedsgrænsen på 50%.

For det andet kan vi konstatere, at diskriminationen af stimuli fra forskellige kategorier ikke afviger væsentligt fra diskriminationen af stimuli fra samme kategori. Hos de danske forsøgspersoner kan vi notere en diskriminations-procent på 69% (mellem stimulus 10 og 11) – kun 11% højere end ved diskrimination af stimuli fra samme kategori. Hos de finske forsøgspersoner udgør den højeste diskrimination af stimuli fra forskellige kategorier kun 59% (mellem stimulus 8 og 9), altså næsten det samme som inden for samme kategori (56%).

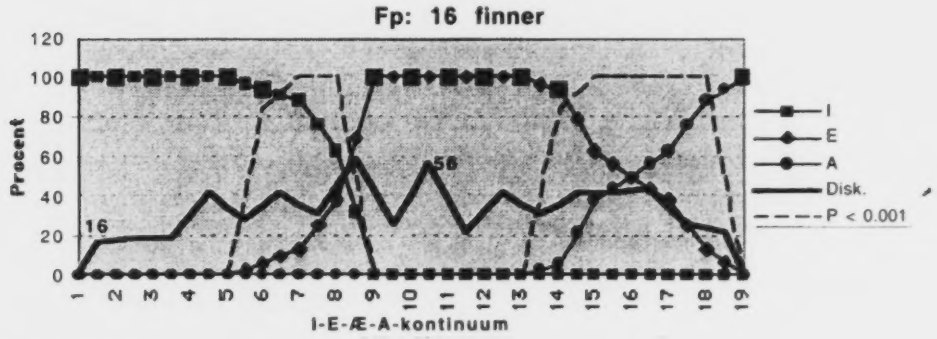
På trods af skarpe grænser mellem hhv. de danske og de finske vokal-kategorier er der altså ikke tale om kategorisk perception, fordi identifikationsfunktionen ikke ledsages af en markant kategorisk effekt.

Statistiske analyser

Hvis vi herefter udelukkende fokuserer på diskriminationen af stimuli, der er signifikante med $P < 0{,}001$, træder et endnu klarere mønster frem. I figur 2.4.8 har jeg sammenholdt de to fp-gruppers identifikationer, identifikationernes signifikansgrænse for $H_0(P < 0{,}001)$ samt diskrimination. Den stiplede signifikanskurve for $P < 0{,}001$ afgrænser hhv. de 8 danske og 13 finske signifikante stimuli, der er markeret med store firkantede koordinater.



114



Figur 2.4.8

Ifølge Haskins-modellen afspejles KP i en meget ringe diskrimination af stimuli fra samme kategori, men relativt god diskrimination af stimuli fra forskellige kategorier.

De 8 stimuli, der fordeler sig over 4 danske kategorier, er ikke alene meget signifikante, men også svære at diskriminere. Som det fremgår af diagrammet finder vi netop den ringeste danske diskrimination mellem signifikante stimuli, nemlig 21% mellem stimulus 14 og 15. Tilsvarende udgør den maksimale diskriminationsprocent af signifikante stimuli kun 33% (mellem stimulus 1 og 2). Omvendt er de signifikante stimuli ikke generelt de sværeste at diskriminere – faktisk har hele 6 par insignifikant kontrasterede stimuli en diskrimination mellem 21 og 33%. Der er altså intet i de danske data, der indikerer en markant forskel i diskriminationen af stimuli fra hhv. forskellige kategorier og samme kategori.

De 13 stimuli, der udgør kernen i de 3 finske kategorier, er ligeledes svære at diskriminere. Igen skal den laveste diskriminationsprocent findes ved diskrimination af signifikante stimuli, nemlig mellem stimulus 1 og 2, der kun er blevet diskrimineret i 16% af tilfældene. Den maksimale diskriminations-procent af signifikante stimuli (stimulus 10 og 11) er derimod helt oppe på 56%, hvilket er den næsthøjeste finske diskriminationsprocent. Heller ikke i de finske data er der indikationer for, at signifikante stimuli generelt er de sværeste at diskriminere – her har hele 7 par insignifikante kontrasterede stimuli en diskrimination mellem 16 og 56%.

Som ovenstående analyser illustrerer, afspejler hverken danske eller finske resultater kategorisk perception af vokallyde. Ganske vist adskilles de fire

danske ('I, 'E, 'Æ og 'A) og tre finske kategorier ('I, 'E og 'A) ved relativt skarpe kategorigrænser, dvs. uden gradueret kategorimedlemskab, men de abrupte overgange i identifikationsfunktionen ledsages ikke af en markant kategorisk effekt, hvorfor identifikations- og diskriminationsadfærden ikke kan betegnes som KP.

Perspektivering af resultater

I det følgende er det min hensigt at sammenholde resultaterne fra de præsenterede identifikations- og diskriminationsserier med hhv. hypoteser, der har haft stor indflydelse på fortolkningen af fænomenet 'kategorisk perception', og mine egne forventninger til udfaldet af de enkelte delforsøg.

Hypotesen om stimulus-asymmetri

Stort set alle beskrivelser af fænomenet 'kategorisk perception' bygger på en generel, dikotomisk fremstilling med perceptionen af hhv. vokaler og konsonanter som to adskilte fænomener. Det hævdes almindeligvis, at lukkelyde kan fremkalde en kategorisk effekt, mens isolerede vokallyde diskrimineres lige godt eller ringe, hvad enten de tilhører samme eller forskellige fonetiske kategorier.

Data fra langt de fleste empiriske undersøgelser viser dog, at situationen er langt mere kompleks: For det første viser forsøg med KP sjældent ideel kategorisk perception, dvs. en diskriminationsadfærd, der udelukkende er betinget af kategoriseringen. Oftest er diskriminationen af stimuli fra samme kategori simpelthen bedre end forventet (jf. side 49). For det andet viser undersøgelser med vokallyde præsenteret under varierede forsøgsbetingelser, at graden af kategorisk effekt kan manipuleres. Der er altså meget, der tyder på, at en dikotomisk beskrivelse af vokal- og konsonantperception ikke er adækvat i forbindelse med forsøg med KP.

Formålet med de nærværende forsøg med KP har først og fremmest været at efterprøve hypotesen om en signifikant forskel mellem perceptionen af vokal- og konsonantlyde. For at fremkalde størst mulig forskel i perceptionen af vokal- og konsonantlyde har jeg valgt at præsentere stimuli under forhold, der postuleres at fremkalde den største effekt, nemlig klusiler i CV-stavelser og isolerede langvokaler.

Det var min forventning, at jeg ville være i stand til at registrere en forskel mellem perceptionen af vokal- og konsonantlyde, dog ikke så stor en forskel som resultaterne afslører.

For det første forventede jeg, at præsentationen af et VOT-kontinuum ville fremkalde et markant skifte i identifikationsfunktionen, en god diskrimination af stimuli fra forskellige kategorier, men en markant ringe diskrimination af stimuli fra samme kategori.

Det forhold, at identifikationen af 10 ud af 14 stimuli hos begge fp-grupper viste sig at være signifikant med $P < 0{,}001$, demonstrerer på bedste vis skiftet i identifikationsfunktionen. De 10 stimuli var dog ikke alene meget signifikante, men også de sværeste at diskriminere. Således udgjorde de signifikante stimuli gruppen af lyde, der havde de absolut laveste diskriminationsprocenter med 28% som det maksimale hos danske forsøgspersoner og 31% som det maksimale hos finske forsøgspersoner. Omvendt havde både de danske og de finske forsøgspersoner noget lettere ved at skelne mellem stimuli fra forskellige kategorier, hvilket afspejler sig i diskriminationsprocenter helt oppe på 88% og 69%. Resultaterne fra forsøget med et VOT-kontinuum opfylder altså mine forventninger til overflod: hverken de danske eller de finske forsøgspersoner kan skelne mellem forskellige stimuli, de identificerer ens.

For det andet forventede jeg, at også præsentationen af et formantkontinuum ville fremkalde et markant skifte i identifikationsfunktionen, men en ringe diskrimination af stimuli – uafhængigt af om de tilhører samme eller forskellige kategorier.

Også ved præsentationen af isolerede vokallyde viste identifikationen sig at være meget signifikant. Hos den finske fp-gruppe var identifikationen af 13 ud af 19 stimuli signifikant med $P < 0{,}001$, hvorimod de danske identifikationer var signifikante på dette niveau i 8 tilfælde. Diskriminationen af stimuli, hvad enten de tilhørte samme eller forskellige kategorier, var derimod temmelig ringe. Hos de danske forsøgspersoner udgjorde den maksimale diskriminationsprocent af signifikante stimuli fra samme kategori faktisk kun 33% , hvorimod den maksimale diskriminationsprocent af stimuli fra samme finske kategori var oppe på 56% – altså lidt over tilfældighedsgrænsen på 50%. Det er dog interessant at notere, at de to eneste

diskriminationstoppe, der overstiger 50% percentilen, findes i samme position hos både danske og finske forsøgspersoner, nemlig mellem stimulus 8 og 9, 10 og 11. Det er desværre ikke muligt at sige – på baggrund af dette materiale – om diskriminationstoppenes position er betinget af ikke-sproglige faktorer, f.eks. intervenerende akustiske variabler, unøjagtigheder i det akustiske kontinuum, fysiologiske faktorer, eller om de udelukkende er opstået ved en tilfældighed.

Jeg kan på baggrund af et solidt empirisk grundlag konkludere, at nærværende forsøg med kategorisk perception afslører en markant forskel i perceptionen af vokal- og konsonantkontinua, en forskel der primært angives ved tilstedeværelsen af en kategorisk effekt. Som tidligere beskrevet understøtter dette resultat motorhypotesen, dvs., at fonetiske kategorier, der adskiller sig ved diskrete artikulatoriske egenskaber, perciperes som absolutte, diskrete størrelser, hvorimod perceptionen er kontinuerlig, når kontinuerlig artikulatorisk variation er mulig.

3.1 Introduktion til dikotisk lytning

Forsøget med kategorisk perception indikerer, at vokal- og konsonantlyde perciperes på forskellig måde. Årsagen til asymmetri i KP af forskellige sproglyde kan naturligvis ikke blotlægges, medmindre man er i stand til at give en beskrivelse af de processer, der er involveret i sprogperception.

Et af de væsentligste spørgsmål er – efter min opfattelse – hvorfor visse stimuli i højere grad end andre kan fremkalde en kategorisk effekt, mens andre slet ikke fremkalder en kategorisk effekt, men snarere tyder på kontinuerlig perception. Resultaterne af en lang række undersøgelser viser, at lukkelyde fremkalder den kraftigste kategoriske effekt, nasaler, lateraler og frikativer en mindre kategorisk effekt, og at vokaler perciperes kontinuert (jf. side 34ff).

De fleste af synteseteorierne tager udgangspunkt i en hypotese om, at sprogperceptionen er kendetegnet ved fonetisk syntese, og at den er neurologisk nært forbundet med sprogartikulation. Således bunder forskellene mellem vokal- og konsonantperceptionen i grundlæggende forskellige processer: lukkelyde består af mere komplekse artikulatoriske kommandoer end vokallyde, der er mere artikulatorisk kontinuerlige, og lukkelydene kræver derfor en speciel sprogprocessor, der skal varetage de komplekse analyser.

Hvis perceptionen af lukkelyde kræver en særlig sprogprocessor, som analyserer de komplekse artikulatoriske kommandoer, forklarer det ligeledes resultater fra et helt andet forsøgsparadigme, nemlig 'dikotisk lytning'.

I forsøg med dikotisk lytning (herefter DL), dvs. dobbeltsidig høre-påvirkning, præsenteres forsøgspersonerne simultant for to stimuli – en i hvert øre – hvorefter de skal identificere én af de pågældende lyde. DL blev opdaget af den engelske psykolog D.E. Broadbent, der som den første konstaterede, at når to simultane lyde bliver præsenteret for de to ører, frembringer det ene øre en hurtigere respons end det andet (Broadbent 1954). Denne effekt er også bedre kendt som højre-ørepræference[1]. HØ-præferencen

[1] Af bekvemmelighedsmæssige grunde vil jeg i det følgende forkorte højre-ørepræference til HØ-præference. Venstre-ørepræference vil af samme grunde blive forkortet til VØ-præference. Forkortelserne korresponderer med de engelske REA (Right Ear Advantage) og LEA (Left Ear Advantage). HØ-præference defineres oftest som: Antal stimulusgenkendelser i

for sproglige stimuli er siden midten af 1950'erne blevet fortolket som udtryk for venstre hjernehalvdels dominans for sprog.

Også i forbindelse med dikotisk præsentation af sproglyde har man registreret en markant forskel mellem lukkelyde og vokallyde, hvilket antyder en lateralisering af den særlige sprogprocessor, der kræves for analysen af lukkelyde, men ikke vokallyde:

1) Lukkelydene fremkalder generelt en relativt stor HØ-præference (Cutting 1974; Crystal & House 1974; Day & Vigorito 1973; Shankweiler & Studdert-Kennedy 1967; Studdert-Kennedy & Shankweiler 1970, etc.).

2) Likvider, frikativer og semivokaler fremkalder generelt en lavere HØ-præference end lukkelydene (Crystal & House 1974; Cutting 1972, 1974; Day & Vigorito 1973; Darwin 1971; Haggard 1971).

3) Isolerede langvokaler fremkalder under normale omstændigheder ikke en HØ-præference, men hævdes endda at kunne fremkalde en VØ-præference (Crystal & House 1974; Cutting 1974; Day & Vigorito 1973; Studdert-Kennedy & Shankweiler 1970, etc.)

Formålet med den følgende undersøgelse er netop at finde ud af, om der eksisterer en sammenhæng mellem resultaterne fra forsøg med KP og DL, dvs. mellem kategorisk og dikotisk effekt.

Hjernens opbygning

Al menneskelig adfærd udgår fra hjernen og kommer til udtryk ved hjælp af muskler og kirtler, såvel som al menneskelig tænkning, sindsbevægelser, hukommelse og viden. Hjernen er ansvarlig for alt det vi er, og alt det vi kan. Menneskets hjerne består af to hjernehalvdele (hemisfærer) – en venstre og en højre. De arbejder forskelligt, modtager forskellige impulser, bearbejder dem forskelligt, men samarbejder også og udveksler information.

Hjernen er et komplekst og indviklet system, der består af et sindrigt opbygget væv. Cortex (hjernebarken) indeholder milliarder af hjerneceller (neuroner), der indgår i et stort antal forbindelser med hinanden. I cortex findes endvidere underafdelinger med meget specialiserede funktioner.

højre øre minus antal stimulusgenkendelser i venstre øre (i %).

Cortex bukker omkring midteraksen i hjernen, således at hver af de to hemisfærer udgør en afrundet struktur, der har forbindelse med den anden halvdel via et stort bundt nervetråde, der kaldes corpus callosum (hjernebjælken). I det indre af hemisfærerne (subcortex) findes ligeledes nervetråde, der forbinder de forskellige dele af cortex med hinanden og med resten af kroppen. Alt dette kaldes tilsammen for cerebrum (storhjernen). Under cerebrum findes cerebellum (lillehjernen), der hovedsagelig har med samordningen af besked til musklerne at gøre (Teyler 1977; Frandsen 1994b).

Venstre hemisfære modtager nerveimpulser fra højre side af kroppen, mens højre hemisfære modtager nerveimpulser fra venstre side af kroppen. Denne form for forbindelse kaldes kontralateral ("modsatsidig"). Hos alle mennesker projiceres stimuli, der præsenteres til venstre for synsfeltets centrum, eller som påvirker venstre hånd eller øre, således direkte til højre hemisfære, og omvendt for stimuli i højre side (Kimura 1964, 1967; Nebes 1977; Frandsen 1994a, etc.). Læsioner af den motoriske cortex fører ligeledes til lammelse af de modsat stillede lemmer, ligesom skader i den postcentrale region af cortex fører til følelsesløshed i den modsatte side af kroppen (Luria 1975). Venstre hemisfære modtager også impulser fra venstre side af kroppen, ligesom højre hemisfære også modtager impulser fra højre side af kroppen. Denne forbindelse kaldes ipsilateral ("sammesidig"). Meddelelserne til de to hemisfærer sker ved hjælp af både ipsi- og kontralaterale forbindelser, men meddelelserne gennem de kontralaterale forbindelser er dog klart de stærkeste.

I det følgende gives en kort – og til tider forsimplet – fremstilling af forskningen i funktionslokalisering, dvs. forskning baseret på en hypotese om, at visse funktioner kan lokaliseres i den menneskelige hjerne. Det skal dog pointeres, at forskningen i funktionslokalisering er kontroversiel – hvad enten den baseres på det 18. århundredes frenologiske studier, studier af unilaterale hjerneskader, eller moderne hjernescanninger, som f.eks. PET- og fMRI-scanninger. Den følgende fremstilling bygger først og fremmest på Segalowitz (1977), Kimura (1961a, 1961b, 1964, 1967), Wittrock (1977). For yderligere diskussion se bl.a. Bates (in press), Paradis (1985, 1995a, 1995b, 1996), Paradis & Gopnik (1994).

I 1861 opdagede den franske anatom og læge Paul Broca, at skader i en bestemt del af hjernen – der nu kaldes for Brocas område – resulterede i en

manglende evne til at frembringe sammenhængende tale. 13 år senere – i 1874 – fandt den tyske psykiater Carl Wernicke endnu et sprogområde i hjernen (Wernickes område). Beskadigelse af dette område resulterer i sammenhængende, men komplet meningsløs tale. Brocas område er således forbundet med evnen til at producere sprog, mens Wernickes tager sig af sprogperceptionen. Brocas område beskæftiger sig først og fremmest med syntaksen, hvorimod Wernickes område styrer de semantiske sider af sproget (Frandsen 1994a; Teyler 1977; Luria 1975). Begge sprogcentre findes i venstre hemisfære hos (næsten) alle højrehåndede og hos en meget stor del af de venstrehåndede.

At venstre hemisfære må være involveret i sprogfunktionen underbygges af flere forskellige forhold. Venstresidige hjernelæsioner fører således langt hyppigere til tab af sprogfunktionen end skader i højre hemisfære. En gruppe børn, der som spæde havde fået fjernet den ene eller den anden hjernehalvdel, blev undersøgt i 10-årsalderen. Når venstre hemisfære manglede, var forståelsen af syntaktiske forhold meget ringe. Når der var pragmatiske træk at basere forståelsen på, gik det bedre (Millar & Whitaker 1983). Hvis venstre hemisfære kortvarigt sættes ud af funktion ved narkose[2] – som det f.eks sker ved hjerneoperationer – resulterer det ligeledes i tab af sprog-funktionen. Dette sker ikke, hvis højre hemisfære bedøves (Wada & Rasmussen 1960). Når verbalt materiale præsenteres i højre eller venstre synsfelt, viser de fleste forsøgspersoner endvidere en højre synsfelt-præference, hvad angår svarenes nøjagtighed og reaktionstid – formodentlig på grund af bedre forbindelser fra højre synsfelt til venstre hjernehalvdel. Ved løsning af verbale opgaver er der ligeledes blevet påvist tegn på større aktivitet i den venstre hemisfære (Morgan et al. 1971; Robbins & McAdam 1974; Wood et al. 1971).

Generelt set foretager venstre hemisfære en sekventiel/seriel behandling af impulserne, mens højre foretager en simultan/parallel. Viser man således en gruppe forsøgspersoner en lang række bogstaver, og måler man deres reaktionstid for at træffe en afgørelse om alle bogstaver er ens, vil de bogstaver, der præsenteres i højre synsfelt (dvs. for venstre hemisfære) blive bearejdet i serie, mens de, der præsenteres i venstre synsfelt (dvs. for højre hemisfære) vil blive bearbejdet parallelt (Nebes 1977) Venstre hemisfære fokuserer på delen frem for helheden, mens højre beskæftiger sig med hel-

[2] Almindeligvis anvendes stoffet sodium amytal (amobarbital), når visse områder af hjernen skal sættes ud af funktion.

heden frem for delen. Venstre hemisfære arbejder begrebsligt-funktionelt, mens højre arbejder visuelt-rumligt. Patienter med skader i højre hemisfære har således meget svært ved at finde vej – selv i vante omgivelser. De kan ikke beskrive velkendte veje, bruge et kort, lægge puslespil eller på anden måde tænke rumligt. Dette sker ikke for mennesker med skader i venstre hemisfære.

Forskellige funktioner kan altså lokaliseres til forskellige områder i de to hemisfærer. De centrale sprogfunktioner synes at befinde sig i venstre hjernehalvdel, men det er mere kompliceret end som så. Ikke alle sprogets aspekter er begrænset hertil, og der synes også at foregå en hel del non-verbalt arbejde i venstre hemisfære. Meget tyder endvidere på, at højre hemisfæres involvering er nødvendig i visse sproglige sammenhænge. Ingen af de to hjernehalvdele kan dog under normale omstændigheder agere uafhængigt af den anden og af resten af kroppen.

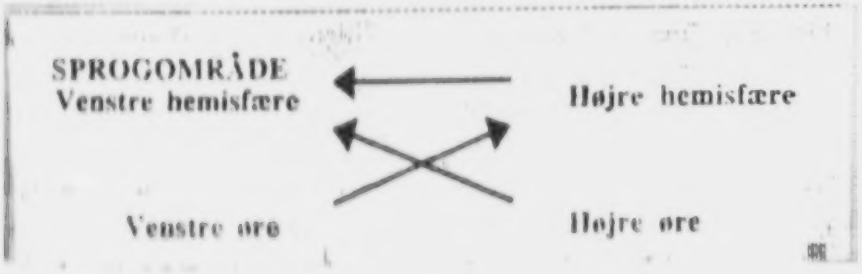
Dikotisk lytning

Ørerne sender impulser til begge hemisfærer. Dette sker ved hjælp af både ipsi- og kontralaterale forbindelser. Med denne viden er testteknikken dikotisk lytning udviklet. DL er dobbeltsidig hørepåvirkning - dvs. det at lytte til forskellige lydmæssige signaler i hvert sit øre (Wittrock 1977).Ved DL hører forsøgspersonerne således simultane stimuli i begge ører. Noget præsenteres for venstre øre, mens noget andet præsenteres for højre.

Metoden til dikotisk lytning blev, som tidligere nævnt, oprindeligt udviklet af den engelske psykolog D.E. Broadbent, der havde opdaget, at det ene øre frembringer en hurtigere respons, når to simultane lyde bliver præsenteret for de to ører (Broadbent 1954). Siden begyndelsen af 1960'erne har metoden hyppigt været benyttet til videre udforskning af asymmetrien mellem menneskets to hjernehalvdele (Kimura 1961a; 1961b; 1967 m.fl.). Tidligere havde undersøgelsen af hjernens funktioner kun baseret sig på observationer af de følger, forskellige hjerneskader førte med sig. Med udviklingen af DL blev det nu muligt at benytte sig af normale forsøgspersoner. Når mennesker med venstrehemisfæredominans for sprog (dvs. størstedelen af menneskeheden) således hører dikotisk præsenteret materiale af verbal karakter, vil højre øre typisk identificere flere stimuli end venstre, hvilket betegnes HØ-præference. Hos mennesker med sprogcenteret repræsenteret i

højre hemisfære, gør det omvendte sig gældende. Her identificerer venstre øre flere stimuli end højre (Kimura 1961b:168).

Kimura (1967) påpeger, at hemisfæreasymmetrien kun kan demonstreres, når der er konkurrence mellem stimuli – dvs. ved dikotisk præsentation. Derimod kan der ikke påvises nogen asymmetri, hvis stimuli præsenteres monauralt[3]. Dette skyldes nervebanernes opbygning. Både de ipsilaterale og de kontralaterale baner sender impulser til hemisfærerne. De kontralaterale baner er imidlertid de stærkeste og dermed i stand til at sætte de ipsilaterale (delvist) ud af kraft. Dette kaldes for okklusionsmekanismen (Kimura 1967:171). Når forskellige stimuli præsenteres for ørerne – sådan som det sker ved DL – vil den stærke impuls fra det kontralaterale øre blokere for den svage impuls fra det ipsilaterale. Reelt modtager hemisfærerne således kun impulserne fra det kontralaterale øre. Når højre øre hos de fleste mennesker identificerer flere sproglige stimuli ved dikotisk præsentation end venstre øre, skyldes det, at højre øre takket være de kontralaterale forbindelsers overlegenhed har bedre adgang til sprogcentrene i venstre hemisfære. Mens højre øre således har direkte forbindelse til venstre hemisfære, skal venstre øres vej til sprogområderne først gå gennem højre hemisfære, før den "krydser over" til venstre side, jf. figur 3.1.1 (fra Krashen 1977:114).



Figur 3.1.1

Dikotisk lytning er oftest blevet brugt til at påvise venstre hjernehalvdels dominans for sprog. Således er HØ-præference blevet påvist mange gange ved dikotisk præsentation af sprogligt materiale. Både meningsfyldte ord, vrøvleord og syntetisk producerede CV-stavelser afslører en HØ-præference (Curry 1968). Donald Dirks (1964) har påvist, at korte ord (f.eks. tal) fremviser en

[3] Ved monaural præsentation afspilles der kun stimuli i et øre ad gangen. Der præsenteres ikke noget for det andet øre.

HØ-præference, ligesom man har fundet eksempler på, at den grammatiske opbygning af sætninger (Zurif & Sait 1969) samt baglæns afspillet tale analyseres bedst af venstre hemisfære. Generelt set afslører dikotisk præsenterede konsonanter en HØ-præference, mens det samme ikke gør sig gældende for vokaler (Day & Vigorito 1973).

Omvendt har man observeret en VØ-præference ved dikotisk præsentation af melodier (Kimura 1964), tonekonturer, miljøstøj (Curry 1967, Spellacy 1970), intonationskonturer (Blumstein & Cooper 1974) samt følelsesmæssige betoninger i sproget (Bryden & Ley 1983).

Den eksperimentelle indgangsvinkel til undersøgelser af hemisfæreasymmetrien har både sine fordele og sine ulemper. Ud over de tidligere omtalte ulemper ved eksperimentelt arbejde (jf. side 5) indebærer forsøg med DL at forsøgspersonerne lytter på en unaturlig måde. Måden man er opmærksom på ved DL er helt forskellig fra den måde, man er opmærksom på i en naturlig situation. I en almindelig kommunikationssituation er opmærksomheden rettet mod forståelsen, mens man ved DL fokuserer på det lydlige udtryksmiddel. Metoden muliggør således undersøgelser af meget specifikke lydlige enkeltfænomener i forhold til hemisfærerne.

De første iagttagelser af, hvor forskellige funktioner kunne lokaliseres i hjernen, blev gjort ud fra observationer af de følger, unilaterale hjerneskader førte med sig. Forskning i hemisfæreasymmetri var således begrænset til at omfatte hjerneskadede personer. Når en person bliver skadet i hjernen, sker der imidlertid også en ændring af denne persons psyke. Hvor meget af en hjerneskadet persons adfærd der skyldes selve hjerneskaden, og hvor meget der skyldes de psykologiske følger af den, synes derfor umuligt at afgøre.

Andre iagttagelser af hjernens funktioner er forsøgt gjort ved hjælp af såkaldte EEG-udsving. EEG (elektroencefalogram) er det billede af elektrisk aktivitet, der kan registreres fra hjernebarken ved brug af overfladeelektroder, der anbringes i hovedbunden (Wittrock 1977). Da EEG afspejler aktiviteten i en stor gruppe neuroner på en gang, er målingen ikke af større værdi ved undersøgelsen af mere specifikke funktioner i hjernen (Teyler 1977).

En af de nyere måder at udforske hjernens funktioner på sker ved hjælp af såkaldte biomagnetiske målinger. Et relativt nyt apparat – PET-scanneren[4] – gør det således muligt at måle de magnetiske kraftfelter i hjernen og derved skabe et ganske nøjagtigt billede af, hvad der foregår i hjernen, når vi foretager os forskellige ting (jf. f.eks. Bottini et al. 1994). Metoden har uden tvivl tiden for sig, men det tekniske udstyr er stadigvæk så kostbart, at kun de færreste har råd til at benytte sig af det. Desuden er der endnu ikke ret mange, der besidder den nødvendige ekspertise til at betjene apparaturet (Hofstätter 1992). De bedste eksisterende metoder til undersøgelse af hemisfæreasymmetri synes derfor at være DL og dens visuelle modsvarighed takistoskopitesten[5].

[4] PET er en forkortelse for 'Positron Emission Tomography'. Ved PET-studier indsprøjtes et glukoselignende stof i forsøgspersoner, hvorefter den neurologiske aktivitet kan måles. Se desuden NeuroReport (8: 1809-1815) for en beskrivelse af undersøgelser med fMRI.

[5] Begge øjne sender synsindtryk fra højre side af synsfeltet til venstre hemisfære og fra venstre halvdel af synsfeltet til højre hemisfære. Omkring synsfeltets midterakse forefindes en række synsceller, der sender impulser til begge hemisfærer. Ønsker man at lateralisere de visuelle stimuli, må man derfor holde en vis afstand til midteraksen. Stimuli må desuden kun fremvises i 150 millisekunder, for at kun den ønskede hemisfære ser noget. Den traditionelle måde at lateralisere visuelle stimuli på er at anvende et takistoskop – dvs. en lysbilledfremviser, der kan indstilles til kun at fremvise stimuli i meget korte perioder (Frandsen 1994a).

3.2 Forsøg med dikotisk lytning

Pilottest og procedure

Forud for dette forsøg har jeg tidligere arbejdet med dikotisk lytning, bl.a i 1994-95, hvor jeg i samarbejde med Jørgen Nørby Jensen udarbejdede en forsøgsrække med titlen 'En undersøgelse af dikotisk lytning som metode til studiet af hemisfæreasymmetri og sprogperception'. I forbindelse med det projekt blev der udført en lang række pilotforsøg. Indledningsvis lavede vi test med både almindelige ord og nonsens-ord. Disse test gav dog ingen konsistente resultater[1]. Herefter udarbejdede vi en pilottest, hvor stimuli ([be], [de], [ge], [pe], [te] og [ke]) én efter én blev præsenteret først for det venstre øre og så for det højre, mens der blev sendt støj (en tone genereret ved en 300 Hz sinuskurve) i det andet. Vi forventede at kunne lateralisere stimuli på denne måde og derved påvise en HØ-præference. Resultaterne af denne test viste imidlertid hverken en højre- eller venstre-ørepræference: hvis der i det første forsøg var venstre-ørepræference, var der i det andet højre-ørepræference; ørepræferencen skiftede så at sige fra forsøg til forsøg. Testen viste, at støj og sproglyde ikke kan fungere som ligeværdige konkurrerende stimuli[2]. Stimuli lateraliseres ikke, og dermed kan der ikke afsløres nogen ørepræference.

Vi udarbejdede herefter en pilottest, hvor de samme stimuli blev præsenteret dikotisk og lateralt (f.eks. blev [be] præsenteret for højre øre, mens [de] sendtes ind i det venstre). Vi bad forsøgspersonerne vurdere, om en given stimulus blev præsenteret eller ej. Svarene blev herefter kategoriseret efter korrekthed. Resultaterne af denne test afslørede en generel HØ-præference – dvs. der var flere korrekte svar ved stimuli præsenteret for højre øre end ved stimuli præsenteret for venstre øre. På baggrund af de erfaringer, som udsprang af de indledende pilottest, udarbejdede vi det endelige forsøg med dikotisk lytning.

Da den beskrevne metode er gennemprøvet med succes, har jeg valgt også at anvende den i dette forsøg med DL. Den eneste afvigelse vil være de stimuli, der indgår i forsøget, samt de to fp-grupper, der ligeledes indgik i forsøget med KP.

[1] Testene med almindelige ord og nonsens-ord blev udført på et meget tidligt tidspunkt i arbejdsprocessen, hvor vi endnu ikke var helt inde i metodikken omkring dikotisk lytning. Dette kan være en af grundene til resultaternes manglende konsistens.
[2] Dette forhold beskrives nærmere s.117f.

Præsentationskriterier

På baggrund af mine egne og andres forsøg inden for forskningsparadigmet 'dikotisk lytning', vil jeg i det følgende opstille en række kriterier, der skal opfyldes i et forsøg med dikotisk præsentation.

1) Konkurrence mellem de præsenterede stimuli er en nødvendig betingelse for at kunne fremprovokere en HØ-præference. Dette skyldes, som tidligere omtalt (s. 112f), at ved præsentation af én stimulus i ét øre varetager de kontralaterale nerveforbindelser kommunikationen med den modsatte hjernehalvdel, hvorimod de ipsilaterale nerveforbindelser kommunikerer repræsentationen[1] til den sammesidige hjernehalvdel. For at kunne fremtvinge en dikotisk effekt skal to lydlige stimuli altså præsenteres konkurrerende, så kun de "stærke" kontralaterale nerveforbindelser sender information og de "svage" ipsilaterale forbindelser sættes ud af funktion. Dette fænomen, at de ipsilaterale nerveforbindelser sættes ud af kraft, betegnes som nævnt hos Kimura "The Occlusion Mechanism" – okklusions-mekanismen (1967:170). Med andre ord: Hvis de ipsilaterale nerve-forbindelser ikke sættes ud af kraft (okklusionsmekanismen), så vil præsentationen af stimuli heller ikke resultere i en HØ-præference. Det forklarer, hvorfor et dikotisk reaktionsmønster ikke kan fremprovokeres i en monaural test.

Okklusionsmekanismen træder i funktion, når to forskellige stimuli præsenteres samtidig (i hvert sit øre). -Men hvad vil det sige, at stimuli præsenteres samtidig? Kimura har studeret dette spørgsmål indgående i en række forsøg. Hun konkluderer følgende:

"When digits are alternated rapidly between ears, but do not actually overlap, there is only a non-significant trend for the right ear to be better" (Kimura 1967:169).

2) Ifølge Kimura o.a., der beskriver metodikken i forsøg med dikotisk lytning, er konkurrence mellem de præsenterede stimuli en nødvendig og tilstrækkelig betingelse for at fremkalde en dikotisk effekt (jf. Kimura 1967). At betingelsen er tilstrækkelig, mener jeg ikke er korrekt. I en pilottest (jf. side 116), hvor de sproglige stimuli [b], [d], [g], [p], [t] og [k] blev præsenteret

[1] Ved 'repræsentation' skal der forstås den pågældende lyds samlede information i hjernen. Denne repræsentation har naturligvis ikke samme form som den akustiske lyd. Præcis hvilken form (og struktur) denne repræsentation har, er et kompliceret spørgsmål, der falder uden for dette forsøgs rammer.

kombineret med en ikke-sproglig distraktionslyd[4] (en 300 Hz sinustone), kunne Jørgen Nørby Jensen og jeg konstatere, at der ingen signifikant ørepræference var at spore (Jensen & Madsen 1995). Der er i øvrigt ikke – så vidt jeg ved – udført vellykkede forsøg med dikotisk effekt, hvor fokuslyden har været af en sproglig art og distraktionslyden af en ikke-sproglig. I tidligere forsøg (Kimura 1964; Curry 1967; Spellacy 1970), hvor ikke-sproglige lyde er blevet præsenteret konkurrerende i kombination med andre ikke-sproglige lyde, er der derimod ofte blevet påvist en signifikant ørepræference (ved disse forsøg en VØ-præference)[5]. Med andre ord er der meget der tyder på, at de konkurrerende lyde skal have samme perceptuelle status, dvs., at repræsentationerne skal have en vis strukturel lighed for at den konkurrerende præsentation vil have en dikotisk effekt. Jeg vil altså konkludere, at konkurrence mellem præsenterede stimuli med samme perceptuelle status er en nødvendig og tilstrækkelig betingelse for at fremkalde en dikotisk effekt.

3) I beskrivelsen af forsøg med dikotisk lytning er det uheldigvis hidtil blevet overset, hvor vigtigt det er, at stimuli har samme perceptuelle status. Det er især uheldigt, fordi denne betingelse giver os nyttige oplysninger om de mentale operationer. I det følgende afsnit vil jeg beskæftige mig med netop dette forhold[6].

Det primære spørgsmål må være: Hvad er forskellen mellem forsøg, hvor stimuli har samme perceptuelle status, og forsøg, hvor stimuli har forskellig perceptuel status? Hvad angår nerveforbindelsernes informationsoverførsel, er der ingen forskel på de to forsøgsopstillinger. I begge tilfælde præsenteres to forskellige stimuli simultant, hvilket aktiverer okklusionsmekanismen. Det er min opfattelse, at nødvendigheden af strukturel lighed ikke skyldes nervebanernes arkitektur, da den ikke forandres af stimulienes karakter. Hvis forskellen var den, at sproglige stimuli blev "opfattet" bedst af venstre hemisfære, og ikke-sproglige stimuli af højre hemisfære, så skulle en test, hvor stimuli havde forskellig status netop give en markant HØ-præference på de sproglige stimuli, og en markant VØ-præference på de ikke-sproglige. Det

[4] Af fremstillingsmæssige årsager har jeg valgt at bruge betegnelserne 'fokuslyd' og 'distraktionslyd', som referencer til henholdsvis de stimuli, forsøgspersonerne skal lytte efter, og de stimuli, der skal virke som distraktion (og effektuere okklusionsmekanismen).
[5] De ikke-sproglige lyde udgør f.eks. melodier, lyde fra omgivelserne som f.eks. hundegøen, smækkende døre, toiletskyllen etc.
[6] Jeg vil gerne pointere, at beskrivelsen er af deduktiv art og uden empirisk belæg.

er da også sund fornuft med vores viden om hemisfærernes funktionsspecialisering, at en ørepræference kun vil opstå, hvis identifikationen af begge stimuli forudsætter én og samme funktion – en funktion, der kun er til stede i den ene hemisfære. Med andre ord forudsætter den dikotiske effekt en hemisfæreasymmetri.

Det er tidligere blevet beskrevet (og illustreret i fig. 3.1.1), at en sproglig stimulus, der præsenteres for venstre øre, først repræsenteres i højre hemisfære, før den repræsenteres i venstre. Den skal billedlig talt "ad en omvej". Derfor vil en sproglig stimulus, der præsenteres for højre øre, blive repræsenteret i venstre hemisfære tidligere end den stimulus, der præsenteres for venstre øre. Og netop den tidslige faktor er afgørende for en dikotisk effekt. Når den sproglige stimulus, der er præsenteret for venstre øre, "ankommer" til venstre hemisfære, så er sprogprocessoren optaget, og stimulusen må derfor stå i venteposition. Denne venteposition er bedre kendt under betegnelsen 'sensorisk hukommelse'. Den sensoriske hukommelse spiller en væsentlig rolle i vores sansning af omverdenen, idet den sammenkæder forskellige sanseoplevelser til en gestalt. Uden den sensoriske hukommelse ville vores sanseoplevelser være ét stort kaos. Men den sensoriske hukommelses dannelse af gestalter betyder også et tab af information: Studiet af subliminal perception har vist, at den sensoriske hukommelse ikke kan rumme mere end 25 billeder pr. sekund. Tilsvarende er der også grænser for, hvor meget lydlig information den sensoriske hukommelse kan fastholde. Det er min opfattelse, at den dikotiske effekt faktisk er et udtryk for tab af information i den sensoriske hukommelse[7]. Tabet af information er størst for de sproglige stimuli, der præsenteres for venstre øre, fordi det er de stimuli, der skal fastholdes i den sensoriske hukommelse. I et forsøg, hvor distraktionslyden er ikke-sproglig, vil den sproglige fokuslyd ikke blive blokeret. Derfor fremkaldes der ingen dikotisk effekt, når de konkurrerende stimuli har forskellig perceptuel status. Ydermere viser Kimuras undersøgelser af overlap i præsentationen af stimuli, at jo kortere tid, der er mellem præsentationen af den lydlige information, desto større tab af information og dermed større ørepræference (Kimura 1967). Eller kort sagt: Jo længere tid mellem præsentationen af to stimuli, desto mindre ørepræference. Jeg vil altså konkludere, at konkurrence mellem

[7] Denne antagelse stemmer da også ganske godt overens med en lang række forsøg, der har haft det formål at undersøge hukommelsens indvirkning på sprogperception, jf. f.eks. Broadbent 1954; Massaro 1970, 1971, 1972, 1975; Massaro & Oden 1980; Pisoni 1972.

simultant præsenterede stimuli med samme perceptuelle status er nødvendige og tilstrækkelige betingelser for at fremkalde en ørepræference. Ørepræferencen vil manifesteres som følge af hemisfæreasymmetri og tab af information i den sensoriske hukommelse.

Hypoteser

Som tidligere beskrevet (jf. side 112ff) afspejler forsøg med DL en lateralisering af de processer, der varetager analysen af visse sproglyde. Meget tyder altså på, at der eksisterer en sammenhæng mellem de lateraliserede processer, der kan fremkalde en dikotisk effekt, og de processer, der i forsøg med KP fremkalder en kategorisk effekt. Det forhold, at den kategoriske og dikotiske effekt er maksimal ved præsentation af klusiler, mindre ved præsentation af likvider, frikativer og semivokaler og helt fraværende ved præsentation af vokaler, kan tolkes på (mindst) to måder:

A. Den generelle funktionsfordeling i hjernen vedrører også analysen af sproglig stimulation, således at venstre hemisfære foretager en sekventiel/seriel behandling af stimuli med fokus på delen frem for helheden, mens højre hemisfære arbejder simultant/parallelt med lyd-gestalter og fokuserer på helheden frem for delen.

B. De specifikt sproglige funktioner befinder sig i venstre hemisfære, hvor en særlig sprogprocessor reducerer en psykoakustisk replika af stimulationen til komplekse artikulatoriske kommandoer. Parallelt med den motoriske analyse af artikulatorisk komplekse sproglyde foretages generelle auditive analyser, der ikke er lokaliseret i kun den ene hemisfære.

Fortolkning A svarer ganske godt til flere af de postulerede psykoakustiske perceptionsmodeller, der antager at sprogperceptionen kan forklares ved generelle auditive principper (jf. Macmillan et al. 1977; Miller et al. 1976; Pastore 1981; Schouten 1980, 1987 etc.). Hvis denne fortolkning er korrekt, skal man i forsøg med DL forvente at finde en HØ-præference for lukkelyde og en VØ-præference for vokaler. Undersøgelser af hjerneskadede personer giver dog ingen indikationer om særlige problemer med identifikation af vokallyde i forbindelse med skader i højre hemisfære eller særlige problemer med identifikation af konsonantlyde (klusiler) i forbindelse med skader i venstre hemisfære. Tilsvarende viser undersøgelser af SLI-patienter[5] heller

[5] SLI er en hyppig anvendt forkortelse for 'Specific Language Impairment', der er et kraftigt

ikke nogen forbindelse mellem perceptionen af hhv. vokal- og konsonantlyde og hemisfæreasymmetrien, som den postuleres i A. Der er altså ingen undersøgelser, som jeg kender til, der taler for tolkning A.

Fortolkning B er i grove træk identisk med Fujisaki & Kawashimas 'dual-process model', der kombinerer motorteoriens antagelse om, at sprog-perceptionen involverer et sæt neurologiske motorkommandoer, med generelle auditive principper for perception af akustisk stimulation (Fujisaki & Kawashima 1969, 1970, 1971; Pisoni 1971, 1972, 1973, 1975; Pisoni & Lazarus 1974; Pisoni & Tash 1974). Hvis tolkning B – og dermed 'dual-process modellen' – er korrekt, skal man i det følgende forsøg med DL forvente disse resultater:

1) De processer, der varetager analysen af klusiler, er lateraliserede, hvilket afspejles i en HØ-præference i forsøg med DL.

2) Analysen af likvider, frikativer og semivokaler er kun delvis afhængig af de lateraliserede processer, hvilket afspejles i en mindre HØ-præference end ved dikotisk præsentation af lukkelyde.

3) De processer, der varetager analysen af vokallyde, er generelle auditive processer og afspejler derfor ikke ørepræference i forsøg med DL.

Stimuli

Jeg anser beskrivelsen af stimuli som en vigtig forudsætning for en rimelig vurdering af forsøgsresultaterne. Når jeg ønsker at understrege dette, så er det fordi jeg mener, at det er meget uheldigt enten helt at udelade eller give en ufuldstændig beskrivelse af stimuli (hvilket ofte er tilfældet, jf. Broadbent 1954; Kimura 1961a, 1961b, 1964; Springer 1971; Curry 1968; Berlin 1973 etc.)[9].

I forsøg med DL er en beskrivelse af stimuli simpelthen en forudsætning for en rimelig tolkning af resultaterne – især i et forsøg som dette, hvor jeg har valgt at fokusere på relationen mellem stimuli og hemisfæreasymmetri. Det er nødvendigt at have et detaljeret kendskab til de præsenterede stimuli for at

voksende forskningsområde (Paradis 1985, 1990; Paradis & Goplik 1994).

[9] Det store antal pilottest, jeg har udført, skyldes bl.a., at det var umuligt at reproducere mange af de forsøg, der er beskrevet i forskellige sammenhænge – beskrivelserne af stimuli var simpelthen så mangelfulde, at der umuligt kunne blive tale om reproduktion.

kunne bedømme, om effekten af præsentationen kan henledes til en hemisfæreasymmetri eller om den blot skyldes forskellen mellem de præsenterede stimuli.

I to af de tre delforsøg med dikotisk præsentation har jeg valgt at anvende de samme stimuli, som blev anvendt i forsøget med kategorisk perception. I det første delforsøg præsenteres forsøgspersonerne for et VOT-kontinuum bestående af 14 stimuli (jf. beskrivelsen side 53ff) – samme kontinuum som i forsøget med KP blev tilskrevet kategorierne 'S, 'T og 'D. I det andet delforsøg præsenteres forsøgspersonerne for et F1-F2-kontinuum bestående af 19 stimuli (jf. beskrivelsen side 59ff) – altså det kontinuum, der i forsøget med KP blev tilskrevet de fire danske kategorier 'I, 'E, 'Æ og 'A eller de tre finske kategorier 'I, 'E og 'A.

I det sidste af de tre delforsøg har jeg valgt at præsentere 5 CV-stavelser, nemlig en lateral, to nasalkonsonanter (herefter 'nasaler') og to frikativer kombineret med et [a]: [la], [ma], [na], [fa] og [va]. Formålet med at præsentere disse stimuli er naturligvis at finde ud af, om lateraler, nasaler og frikativer resulterer i en dikotisk effekt, dvs. en HØ- eller VØ-præference.

Som første trin i udarbejdelsen af de 5 stimuli indtalte jeg stavelserne [la], [ma], [na], [fa] og [va] på en bærbar pc ved hjælp af dens indbyggede mikrofon[10]. De 5 stimuli blev indtalt direkte på én 16-bit lydfil (44.100 Hz). Jeg valgte at indtale CV-stavelserne så monotont som muligt med meget korte mellemrum på én lydfil for at undgå alt for stor variation. Dernæst gik jeg i gang med at tilpasse den alt for lange lydfil til 5 separate lydfiler med de digitaliserede signaler. I alle andre henseender blev de 5 CV-stavelser [la], [ma], [na], [fa] og [va] bearbejdet på samme måde som de andre stimuli (jf. side 52ff).

Alle de stimuli, jeg har beskrevet ovenfor, udgør fokuslyden i den dikotiske præsentation. Men for at kunne fremkalde en dikotisk effekt, skal man simultant præsentere to stimuli, en fokuslyd og en distraktionslyd. Som distraktionslyd har jeg valgt én stimulus, der skal udgøre distraktionslyd i alle

[10] Oprindelig var det kun min intention at anvende indtalingen i en pilottest, men da lydkvaliteten viste sig at være rimelig høj og forsøgsresultaterne konsistente, så jeg ingen grund til at ændre på stimuli. Desuden skal det tilføjes, at det akustiske signal, vi modtager i en almindelig dagligdags samtale, ofte er af langt ringere kvalitet, men sjældent giver perceptuelle problemer.

forsøgene, nemlig [kɑ:]. Fordelen ved denne distraktionslyd er, at den kan opfattes på flere måder – som et dansk ord (sg. af substantivet 'karre') og som et engelsk ord (sg. af substantivet 'car') – hvilket er en nødvendighed, når man arbejder med flere sproggrupper (her danske og finske forsøgspersoner). I forsøget med DL består alle stimuli af en fokuslyd og en distraktionslyd, der er indspillet på en to-spors lydfil, således at fokuslyden er på det ene spor og distraktions-lyden på det andet spor.

Procedure

Proceduren i afviklingen af forsøget med DL svarer stort set til forsøget med KP. For det første blev forsøget med DL afviklet ved hjælp af præsentations- og registreringsprogrammet DikoTest 3.0, der var installeret på en bærbar pc. De anvendte præsentationsserier var alle bygget op efter det samme princip: først defineres en to-sporet *.wav-lydfil, som DikoTest kan afspille i medieafspilleren. Dernæst præsenteres den tilhørende tekst på skærmen:

Identifikationsøvelser med dikotisk præsentation[11]:

Konsonanter:	'HØRER DU ET D, S ELLER T?	SVAR MED ÉN TAST. #XX'
[la-ma-na-fa-va]:	'HØRER DU ET F, L, M, N ELLER V?	SVAR MED ÉN TAST. #XX'
Vokaler:	'HØRER DU ET A, E, I ELLER Æ?	SVAR MED ÉN TAST. #XX'

Identifikationsserierne for det akustiske S-T-D-kontinuum er identiske med de anvendte serier i forsøget med KP, dvs., at hver identifikationsserie består af 28 præsentationer.

Identifikationsserierne for det syntetiske vokalkontinuum er ligeledes identiske med de identifikationsserier, der blev anvendt i forsøget med KP, dvs., at hver identifikationsserie består af 19 præsentationer.

For de nye stimuli [la], [ma], [na], [fa] og [va] blev der fremstillet en identifikationsserie bestående af 20 præsentationer, dvs., at hver af de 5 stimuli blev præsenteret 2 gange for HØ og 2 gange for VØ.

Også i forbindelse med dikotisk præsentation blev hver præsentationsserie indledt af en "introduktionsrunde" bestående af 6 identifikationsøvelser, der

[11] '#xx' repræsenterer et præsentationsnummer, som forsøgspersonerne skulle notere, hvis deres respons via tastaturet var forkert, dvs. utilsigtet. Igen blev responsmulighederne opstillet i alfabetisk rækkefølge.

ikke er medregnet som forsøgsresultater. De 4 basisserier for S-T-D-identifikation, L-M-N-F-V-identifikation og I-E-Æ-A-identifikation betegnes herefter 'C-Iden-Basis.XLS', 'X-Iden-Basis.XLS' og 'V-Iden-Basis.XLS'.

Forsøgsforhold og forsøgspersoner

Forsøget med DL fandt i alle tilfælde sted i forlængelse af forsøget med KP, og jeg fandt det ikke nødvendigt at oplyse forsøgspersonerne om, at der faktisk var tale om to særskilte forsøg.

I forsøget med DL deltog de samme forsøgspersoner som i forsøget med KP, dvs., i alt 57 personer fordelt på 41 danskere og 16 finner, 42 kvinder og 15 mænd i alderen 12 - 61 år. Ingen af forsøgspersonerne medvirkede flere gange.

Det var ikke alle forsøgspersoner, der responderede på samtlige 3 testserier – enkelte deltagere valgte at springe fra efter kun at have gennemgået én af de to hovedkategorier, konsonantidentifikation, vokalidentifikation eller identifikation af [la-ma-na-fa-va][12]. Desuden valgte jeg at præsentere [la], [ma], [na], [fa] og [va] for de danske forsøgspersoner, men ikke for de finske, hvilket skyldes, at disse stimuli ikke alle korresponderer med finske sproglyde.

I registreringsskemaet har hver enkel forsøgsperson oplyst, om vedkommende er højrehåndet eller venstrehåndet. På baggrund af disse oplysninger var det min hensigt at inddele den samlede fp-population i to subkategorier, nemlig i én bestående af højrehåndede og en anden af venstrehåndede forsøgspersoner. Som tidligere omtalt (side 110) er det dog kun meget få af de venstrehåndede, der rent faktisk har det centrale sprogcenter placeret i højre hemisfære. Da ingen af forsøgspersonerne udviste en adfærd, der kunne tolkes som en lateralisering omvendt af majoritetens, besluttede jeg at behandle forsøgsresultater fra de to subkategorier under ét.

[12] Jeg forsøgte lige som tidligere at fordele det samlede antal forsøgspersoner over de 5 varianter af hver præsentationsserie, jf. figur 3.2.1.

Danske fp:

	Køn	Alder	Præs.serie	VOT-kontinuum	Formant-kontinuum	[la-ma-na-fa-va]
AAO	♂	57	3	Ja	Ja	Nej
BDL	♂	45	4	Ja	Ja	Ja
BRB	♂	27	2	Ja	Ja	Ja
CHW	♀	29	5	Ja	Ja	Ja
COM	♀	23	1	Ja	Ja	Ja
DOB	♀	34	1	Ja	Ja	Ja
FLJ	♂	47	3	Ja	Ja	Ja
GIM	♀	29	4	Ja	Nej	Ja
GNE	♀	27	4	Ja	Ja	Ja
HAF	♀	46	4	Ja	Ja	Ja
HEP	♀	49	2	Ja	Ja	Ja
HES	♂	32	2	Ja	Ja	Ja
IKM	♀	61	5	Ja	Ja	Ja
IMN	♀	30	1	Ja	Ja	Ja
JAO	♀	25	1	Ja	Ja	Ja
JEK	♂	27	1	Ja	Ja	Ja
JNJ	♂	26	3	Ja	Ja	Ja
JOJ	♀	24	2	Ja	Ja	Ja
JTR	♂	35	2	Ja	Ja	Ja
KAM	♀	36	5	Nej	Ja	Ja
KAO	♀	52	4	Ja	Ja	Nej
KIG	♀	52	5	Ja	Ja	Ja
KRR	♀	24	3	Ja	Ja	Ja
LDL	♀	41	5	Ja	Ja	Ja
LER	♀	24	1	Nej	Ja	Nej
LIL	♀	14	1	Ja	Ja	Ja
LOK	♀	25	1	Ja	Ja	Ja
MAN	♀	20	2	Ja	Ja	Ja
MCK	♀	24	3	Ja	Ja	Ja
MDL	♀	12	2	Ja	Nej	Ja
MOH	♂	26	1	Ja	Ja	Nej
MOS	♂	42	5	Ja	Ja	Ja
OLG	♀	45	1	Ja	Ja	Ja

BEST COPY AVAILABLE

PES	♂	28	1	Ja	Nej	Ja
SAH	♀	24	1	Ja	Ja	Ja
SMF	♀	24	1	Ja	Ja	Ja
SOD	♀	29	1	Ja	Ja	Ja
SUM	♀	36	1	Ja	Ja	Ja
SOH	♂	26	1	Ja	Ja	Ja
TES	♀	31	1	Ja	Ja	Ja
TRT	♀	26	1	Ja	Ja	Ja

Uvske fp:

	Køn	Alder	Præsentation	VOT-kontinuum	Formant kontinuum	[ba-ma-na-fa-va]
ANK	♀	24	1	Ja	Ja	Nej
ELK	♀	21	1	Ja	Ja	Nej
JAH	♂	25	1	Ja	Ja	Nej
JON	♀	23	1	Ja	Ja	Nej
KAS	♀	21	2	Ja	Ja	Nej
KAU	♀	23	1	Ja	Ja	Nej
KAV	♂	44	2	Ja	Ja	Nej
LES	♀	21	1	Ja	Ja	Nej
LOP	♀	22	1	Ja	Ja	Nej
LOR	♀	22	1	Ja	Ja	Nej
MAL	♀	22	1	Ja	Ja	Nej
MEO	♀	23	1	Ja	Ja	Nej
MIK	♀	21	1	Ja	Ja	Nej
OUS	♀	23	1	Ja	Ja	Nej
REK	♀	21	1	Ja	Ja	Nej
SHA	♀	22	2	Ja	Ja	Nej

Figur 3.2.1

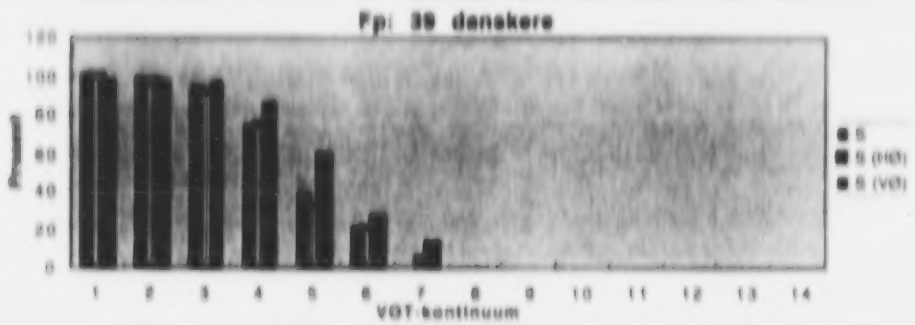
BEST COPY AVAILABLE

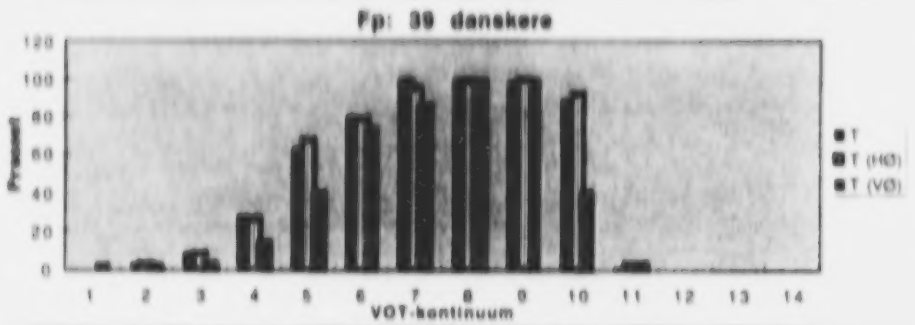
3.3 Resultater fra forsøg med dikotisk lytning: et S-T-D-kontinuum

Det er min hensigt i det følgende at præsentere resultater fra forsøg med dikotisk præsentation af et VOT-kontinuum, et formantkontinuum og af [la], [ma], [na], [fa], [va]. Desuden vil jeg analysere de præsenterede resultater med henblik på de opstillede hypoteser om hemisfæreasymmetri i kapitel 3.2.

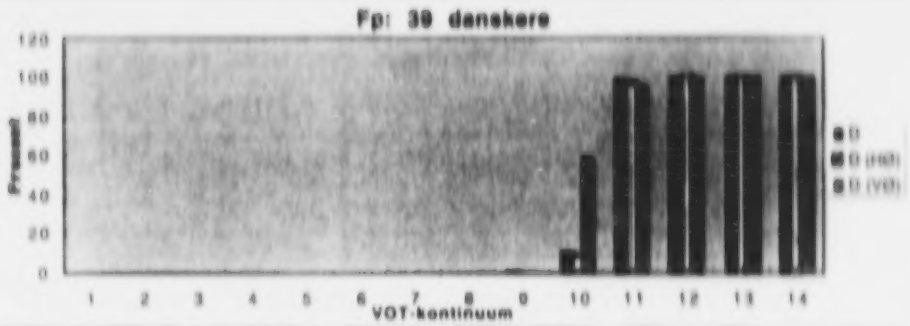
Resultaterne stammer fra de to tidligere omtalte fp-grupper bestående af 41 danskere og 16 finner. Der vil ikke blive skelnet mellem højrehåndede og venstrehåndede, da alle forsøgspersonerne øjensynlig stemmer med normen, dvs., at de har den centrale sprogprocessor i venstre hemisfære.

De første resultater, jeg vil præsentere, er de danske og de finske forsøgspersoners identifikation af 14 stimuli som enten 'dale', 'sale' eller 'tale'. Alle 14 stimuli blev præsenteret kombineret med en konkurrerende distraktionslyd.





BEST COPY AVAILABLE

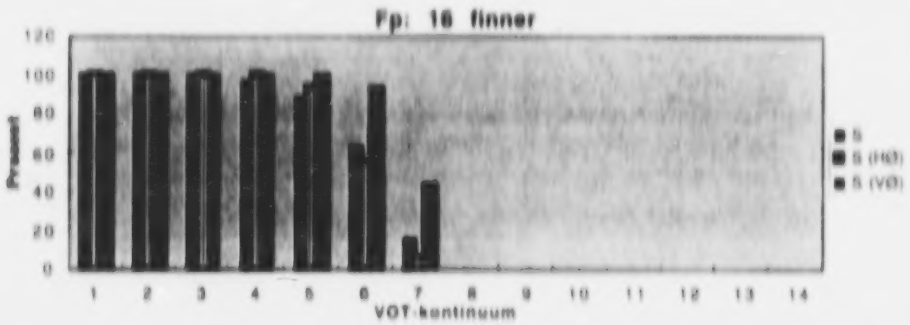


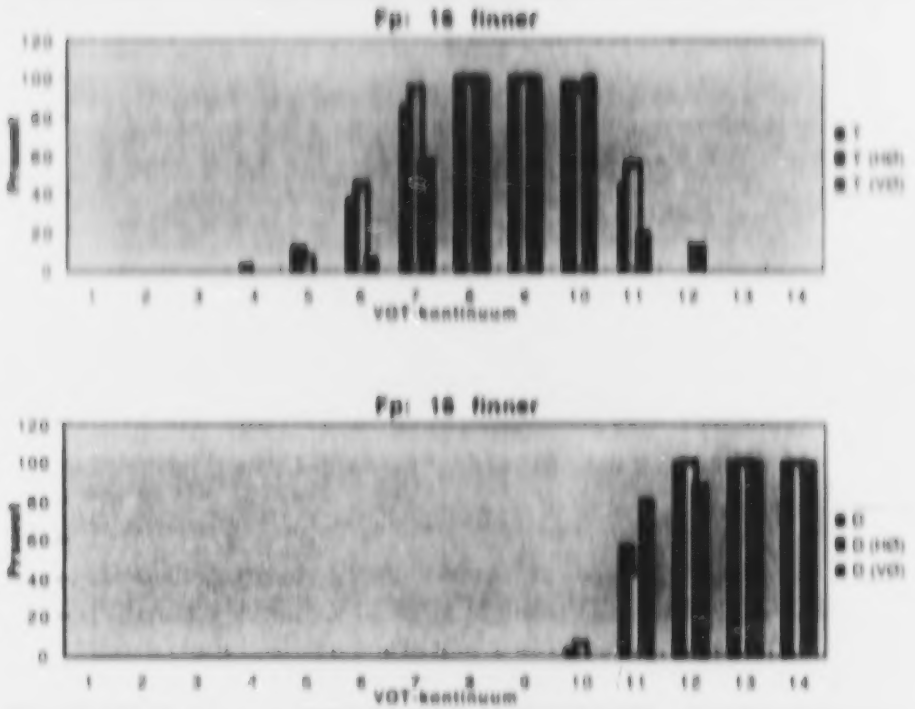
Figur 3.3.1

Gruppen af danske forsøgspersoner bestod af 39 personer, hvilket betyder, at procentfordelingen i figur 3.3.1 er baseret på 546 præsentationer og identifikationer (39 · 14).

Som det tydeligt fremgår af modellen, har forsøgspersonerne kunnet identificere alle stimuli som tilhørende én af de tre kategorier, både ved binaural og dikotisk præsentation. Endvidere fremgår det, at identifikationer ved dikotisk præsentation med fokuslyden i HØ og distraktionslyden i VØ i høj grad stemmer overens med identifikationer ved binaural præsentation. Derimod afviger identifikationer ved dikotisk præsentation med fokuslyden i VØ på visse punkter fra identifikationer ved binaural præsentation. Kort sagt afspejler modellen en generel HØ-præference.

I histogrammet over de finske forsøgspersoners identifikation af samme VOT-kontinuum, præsenteret dikotisk og binauralt, finder vi samme generelle HØ-præference, jf. figur 3.3.2.





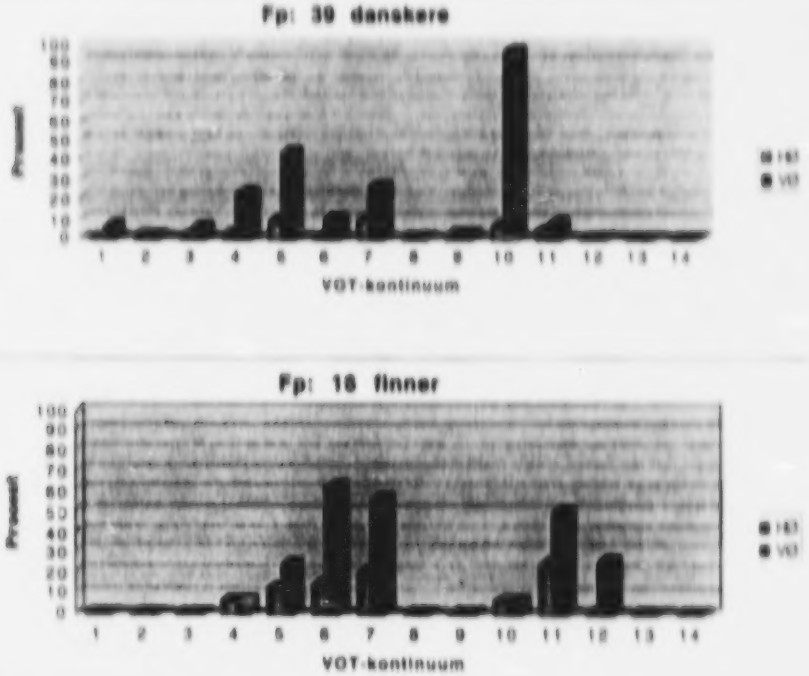
Figur 3.3.2

Gruppen af finske forsøgspersoner bestod af 16 personer, hvilket betyder, at procentfordelingen i figur 3.3.2 bygger på 224 præsentationer og identifikationer (16 · 14).

Også i figur 3.3.2 ser vi, identifikationer ved binaural præsentation i højere grad stemmer overens med identifikationer ved dikotisk præsentation med fokuslyden i HØ end ved dikotisk præsentation med fokuslyden i VØ. De finske resultater afspejler altså også en generel HØ-præference.

Ser vi nærmere på, hvordan de to grupper har tilskrevet kategorier til det akustiske kontinuum – præsenteret hhv. dikotisk og binauralt – træder nogle forskelle frem. Hvis vi fortolker forskellen mellem dikotisk præsenterede stimuli og binauralt præsenterede stimuli som udtryk for hemisfære-asymmetri, er det muligt at fremhæve HØ-præferencen. I figur 3.3.3 fokuserer jeg udelukkende på den akkumulerede forskel (i procent) mellem identifikationer ved binaural præsentation og ved dikotisk præsentation med

fokuslyden i HØ, og den akkumulerede forskel mellem identifikationer ved *binaural præsentation* og ved dikotisk præsentation med fokuslyden i VØ.



Figur 3.3.3

Ved at fremhæve den akkumulerede forskel mellem identifikationer ved "normal" præsentation, dvs. binaural præsentation, og dikotisk præsentation, kan vi illustrere to væsentlige forhold:

1) For det første er HØ-præferencen meget tydelig. I forhold til identifikationer ved binaural præsentation afviger dikotisk præsentation af fokuslyde i HØ mindre – eller lige så lidt, men aldrig mere – end dikotisk præsentation af fokuslyde i VØ.

2) For det andet er HØ-præferencen tydeligvis større ved de sprogspecifikke kategorigrænser. De tidligere beskrevne identifikationsøvelser blotlagde forskelle mellem de danske og de finske kategorier og kategorigrænser: forsøg med danske forsøgspersoner afslørede en kategorigrænse mellem [s] og [t] ved stimulus 4 og 5, samt en kategorigrænse mellem [t] og [d] ved

BEST COPY AVAILABLE

stimulus 10 og 11. I forsøget med DL finder vi netop den højeste H0-præference ved disse kategorigrænser. Identifikationsøvelsen med de finske forsøgspersoner afslørede ligeledes en kategorigrænse mellem [s] og [t], men hos de finske forsøgspersoner mellem stimulus 6 og 7, og en kategorigrænse mellem [t] og [d] ved stimulus 11. Også hos de finske forsøgspersoner finder vi altså en korrelation mellem H0-præference og kategorigrænser.

Årsagen til at H0-præferencen korrelerer med kategoritilskrivningen er sandsynligvis, at stimuli, der under alle omstændigheder er svære at identificere, lettere kan maskeres af en distraktionslyd, end stimuli, der generelt er lette at identificere. Med andre ord kan en høj identifikationsprocent udligne øredifferencen. Denne fortolkning er i øvrigt helt i overensstemmelse med betingelserne for perceptuel overstimulation i forsøg med DL (jf. beskrivelsen af okklusionsmekanismen side 116).

I forsøget på at belyse ovenstående problematik har jeg beregnet sandsynligheden for den konkrete kategoritilskrivning. Til dette formål har jeg opstillet en 'nul-hypotese', H_0, der skal biotlægge sandsynligheden for, at en given stimulus ikke identificeres (jf. beskrivelsen af H_0 side 77). Som det fremgår af P-værdierne i figur 3.3.4, er de stimuli, der har fremkaldt den største H0-præference, samtidig de mindst signifikante stimuli.

Population: danske forsøgspersoners identifikationer
Kategori: S

Stimulus	n (H0)	%	X	S	Z	P	n (V10)	%	X	S	Z	P
1	39	100	1.0	0	0	***	38	97	0.974	0.160	6.09	***
2	38	97	0.974	0.160	6.09	***	38	97	0.974	0.160	6.09	***
3	36	92	0.923	0.270	3.42	***	37	95	0.949	0.223	4.26	***
4	29	74	0.744	0.442	1.683	0.046*	33	85	0.846	0.366	2.31	0.01*
5	13	33	0.333	0.478	0.697	0.242	23	59	0.590	0.498	1.18	0.119
6	8	21	0.205	0.409	0.501	0.308	10	26	0.256	0.442	0.58	0.281
7	2	5	0.051	0.223	0.229	0.409	5	13	0.128	0.339	0.38	0.352
8	0	0	0	0	0	1.0	0	0	0	0	0	0
9	0	0	0	0	0	1.0	0	0	0	0	0	0
10	0	0	0	0	0	1.0	0	0	0	0	0	0
11	0	0	0	0	0	1.0	0	0	0	0	0	0
12	0	0	0	0	0	1.0	0	0	0	0	0	0

142

13	0	0	0	0	0	1.0	0	0	0	0	0	0
14	0	0	0	0	0	1.0	0	0	0	0	0	0

Kategori: I

Stimulus	n (H&I)	%	X	S	Z	P	n (V&I)	%	X	S	Z	P
1	0	0	0	0	0	1.0	38	3	0.026	0.160	-0.16	0.436
2	1	3	0.026	0.160	-0.16	0.436	38	3	0.026	0.160	-0.16	0.436
3	3	8	0.077	0.270	-0.29	0.386	37	5	0.051	0.223	-0.23	0.409
4	10	26	0.256	0.442	-0.58	0.281	33	15	0.154	0.336	-0.42	0.337
5	26	67	0.667	0.478	-1.4	0.081	23	41	0.410	0.498	-0.82	0.206
6	31	79	0.795	0.409	-1.94	0.026*	10	74	0.744	0.442	-1.68	0.046*
7	37	95	0.949	0.223	-4.26	***	3	87	0.872	0.339	-2.57	0.005*
8	39	100	1.0	0	0	***	0	100	1.0	0	0	***
9	39	100	1.0	0	0	***	0	100	1.0	0	0	***
10	36	92	0.923	0.270	-3.42	***	0	41	0.410	0.498	-0.82	0.206
11	1	3	0.026	0.160	-0.16	0.436	0	5	0.051	0.223	-0.22	0.413
12	0	0	0	0	0	1.0	0	0	0	0	0	1.0
13	0	0	0	0	0	1.0	0	0	0	0	0	1.0
14	0	0	0	0	0	1.0	0	0	0	0	0	1.0

Kategori: II

Stimulus	n (H&I)	%	X	S	Z	P	n (V&I)	%	X	S	Z	P
1	0	0	0	0	0	1.0	0	0	0	0	0	1.0
2	0	0	0	0	0	1.0	0	0	0	0	0	1.0
3	0	0	0	0	0	1.0	0	0	0	0	0	1.0
4	0	0	0	0	0	1.0	0	0	0	0	0	1.0
5	0	0	0	0	0	1.0	0	0	0	0	0	1.0
6	0	0	0	0	0	1.0	0	0	0	0	0	1.0
7	0	0	0	0	0	1.0	0	0	0	0	0	1.0
8	0	0	0	0	0	1.0	0	0	0	0	0	1.0
9	0	0	0	0	0	1.0	0	0	0	0	0	1.0
10	1	3	0.077	0.270	-0.29	0.386	23	59	0.590	0.498	-1.18	0.119
11	38	97	0.974	0.160	-6.09	***	37	95	0.949	0.223	-4.26	***
12	39	100	1.0	0	0	***	34	100	1.0	0	0	***
13	39	100	1.0	0	0	***	39	100	1.0	0	0	***
14	39	100	1.0	0	0	***	39	100	1.0	0	0	***

BEST COPY AVAILABLE

Population: Finske forsøgspersoners identifikationer

Kategori: S

Stimulus	n (116)	%	λ	S	Z	P	n (V6)	%	λ	S	Z	P
1	16	100	1.0	0	0	***	16	100	1.0	0	0	***
2	16	100	1.0	0	0	***	16	100	1.0	0	0	***
3	16	100	1.0	0	0	***	16	100	1.0	0	0	***
4	16	100	1.0	0	0	***	16	100	1.0	0	0	***
5	15	94	0.938	0,250	-3.75	***	16	100	1.0	0	0	***
6	9	56	0.563	0,389	-1.45	0,074	15	94	0.938	0,250	-3.75	***
7	1	6	0.063	0,250	-0.25	0,401	7	44	0.438	0,389	-1.13	0,129
8	0	0	0	0	0	1,0	0	0	0	0	0	1,0
9	0	0	0	0	0	1,0	0	0	0	0	0	1,0
10	0	0	0	0	0	1,0	0	0	0	0	0	1,0
11	0	0	0	0	0	1,0	0	0	0	0	0	1,0
12	0	0	0	0	0	1,0	0	0	0	0	0	1,0
13	0	0	0	0	0	1,0	0	0	0	0	0	1,0
14	0	0	0	0	0	1,0	0	0	0	0	0	1,0

Kategori: T

Stimulus	n (116)	%	λ	S	Z	P	n (V6)	%	λ	S	Z	P
1	0	0	0	0	0	1,0	0	0	0	0	0	1,0
2	0	0	0	0	0	1,0	0	0	0	0	0	1,0
3	0	0	0	0	0	1,0	0	0	0	0	0	1,0
4	0	0	0	0	0	1,0	0	0	0	0	0	1,0
5	1	6	0.063	0,250	-0.25	0,401	0	0	0	0	0	1,0
6	7	44	0.438	0,389	-1.13	0,129	1	6	0.063	0,250	-0.25	0,401
7	15	94	0.938	0,250	-3.75	***	9	56	0.563	0,389	-1.45	0,074
8	16	100	1.0	0	0	***	16	100	1.0	0	0	***
9	16	100	1.0	0	0	***	16	100	1.0	0	0	***
10	15	94	0.938	0,250	-3.75	***	16	100	1.0	0	0	***
11	9	56	0.563	0,389	-1.45	0,074	3	19	0.188	0,403	-0.47	0,319
12	0	0	0	0	0	1,0	2	13	0.125	0,342	-0.37	0,356
13	0	0	0	0	0	1,0	0	0	0	0	0	1,0
14	0	0	0	0	0	1,0	0	0	0	0	0	1,0

BEST COPY AVAILABLE

Kategori: D

Stimulus	n (H63)	%	x̄	S	Z	P	n (V63)	%	x̄	S	Z	P
1	0	0	0	0	0	1,0	0	0	0	0	0	1,0
2	0	0	0	0	0	1,0	0	0	0	0	0	1,0
3	0	0	0	0	0	1,0	0	0	0	0	0	1,0
4	0	0	0	0	0	1,0	0	0	0	0	0	1,0
5	0	0	0	0	0	1,0	0	0	0	0	0	1,0
6	0	0	0	0	0	1,0	0	0	0	0	0	1,0
7	0	0	0	0	0	1,0	0	0	0	0	0	1,0
8	0	0	0	0	0	1,0	0	0	0	0	0	1,0
9	0	0	0	0	0	1,0	0	0	0	0	0	1,0
10	1	6	0,063	0,250	-0,25	0,401	0	0	0	0	0	1,0
11	7	44	0,438	0,389	1,13	0,129	13	81	0,813	0,403	2,02	0,022*
12	16	100	1,0	0	0	***	14	88	0,875	0,342	2,56	0,005*
13	16	100	1,0	0	0	***	16	100	1,0	0	0	***
14	16	100	1,0	0	0	***	16	100	1,0	0	0	***

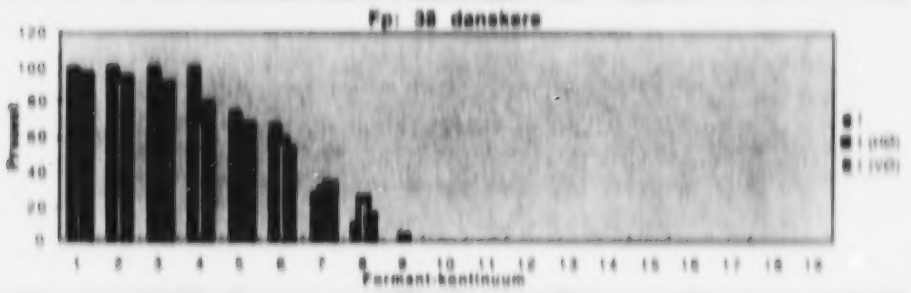
n (H63)/n (V63)	Observationsværdi ved præsentation af fokuslyd for H63/V63
%:	Procent
x̄:	Den aritmetiske middelværdi
S:	Standardafvigelsen
Z:	Observationsværdiens afvigelse fra den aritmetiske middelværdi udtrykt i standardafvigelsesenheder
P:	Sandsynlighed for H_0
*:	$P < 0,05$
***:	$P < 0,001$

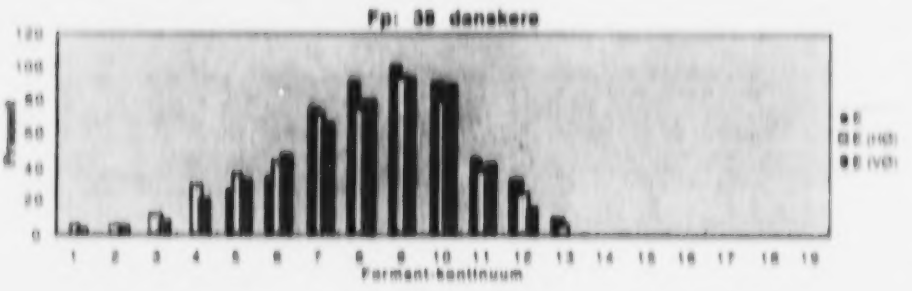
BEST COPY AVAILABLE

3.4 Resultater fra forsøg med dikotisk lytning: et I-E-/Æ-A-kontinuum

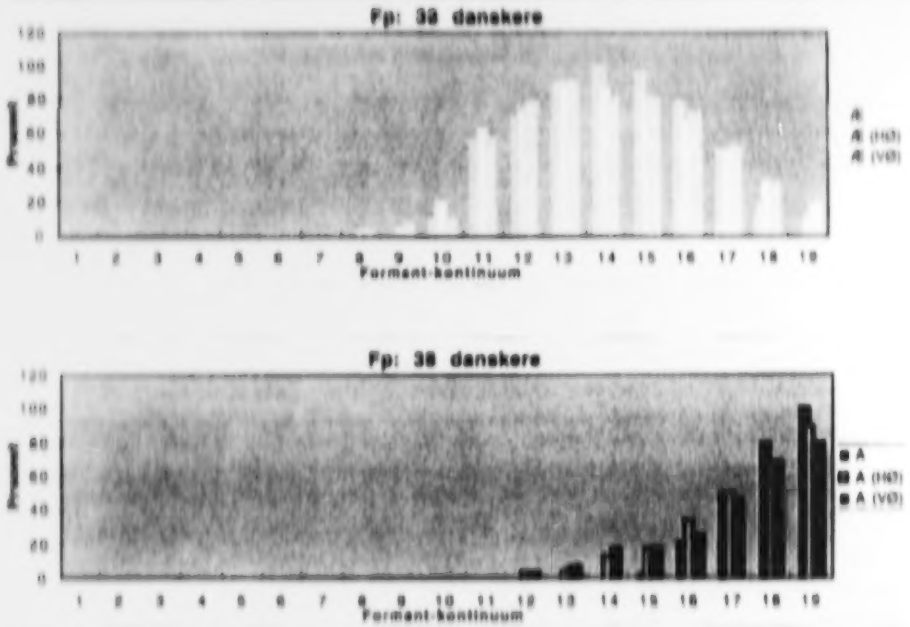
I det foregående kapitel så vi, at perceptionen af dikotisk præsenterede stimuli fra et VOT-kontinuum afspejler en HØ-præference, hvilket kan tolkes som udtryk for hemisfæreasymmetri både hos danske og finske forsøgspersoner. Hvis hypotesen om en lateraliseret sprogprocessor, der involverer neurologiske motorkommandoer, udelukkende vedrører artikulatorisk komplekse sproglyde, kan vi forvente, at perceptionen af isolerede vokallyde ikke afspejler en HØ-præference. Er de processer, der varetager analysen af vokallyde, generelle auditive processer, kan vi endvidere forvente, at en dikotisk præsentation ikke fremkalder nogen ørepræference overhovedet.

I det følgende er det min hensigt at præsentere resultater fra forsøg med dikotisk præsentation af et syntetisk produceret formantkontinuum. Resultaterne stammer fra to fp-grupper bestående af hhv. 38 danskere og 16 finner, der alle har identificeret 19 stimuli som enten 'A', 'E', 'I', 'Æ' eller 'A', 'E', 'I'. Hver stimulus bestod af en fokuslyd – vokallyden – og en distraktionslyd.



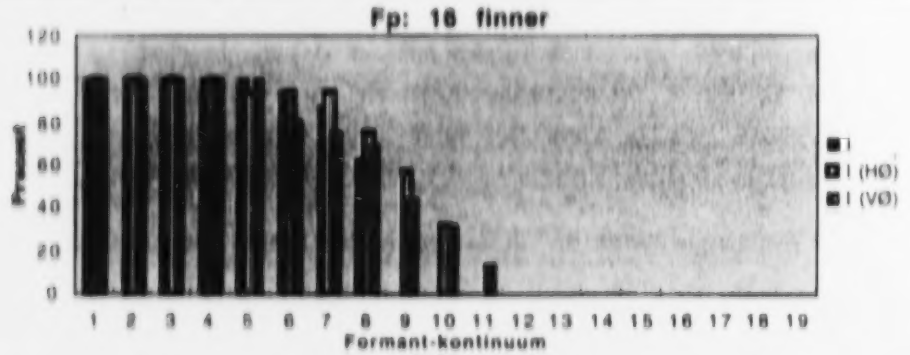


BEST COPY AVAILABLE

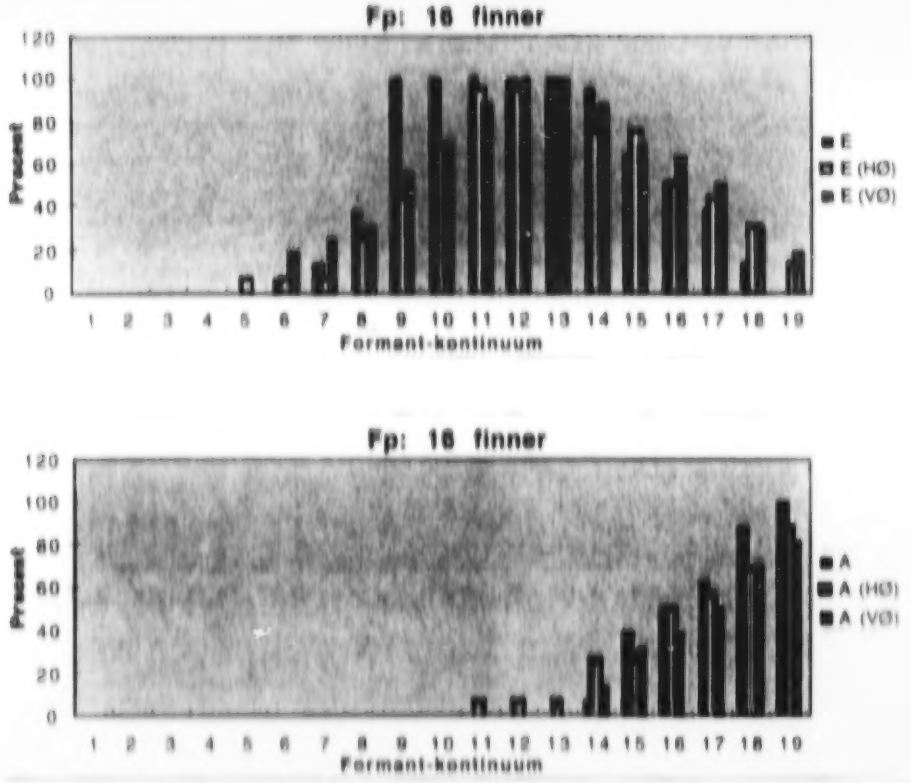


Figur 3.4.1

Figur 3.4.1 illustrerer de danske forsøgspersoners identifikationer af de 19 stimuli. Procentfordelingen bygger på 722 præsentationer og identifikationer (38 · 19). Som det fremgår af histogrammet er der ingen generel tendens til HØ-præference. Derimod lader det til, at dikotisk præsentation af vokallyde fremkalder en diffus fordeling af identifikationer – hvor præsentationen af nogle stimuli fremkalder en HØ-præference, foranlediger andre en VØ-præference.



147



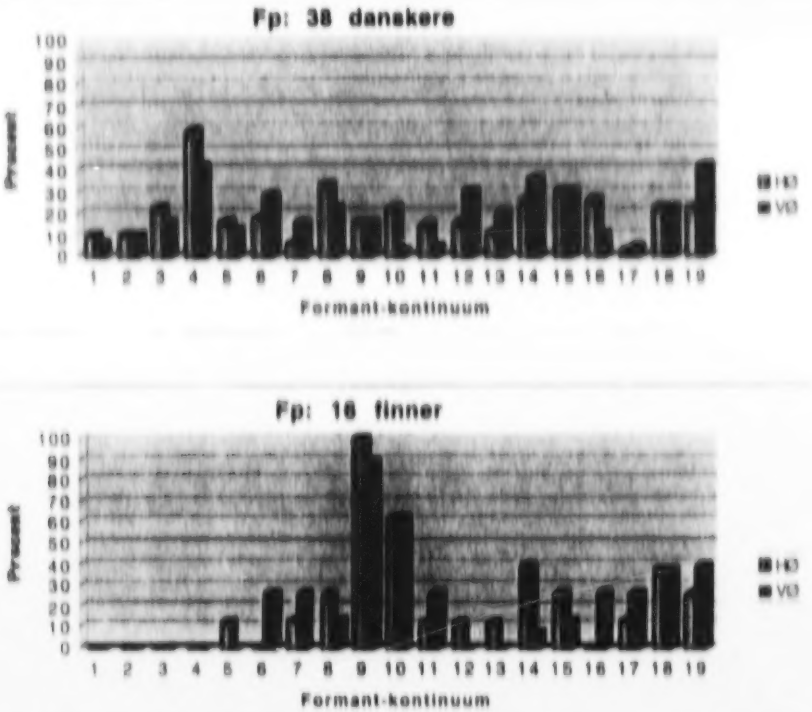
Figur 3.4.2

I modellen over de finske forsøgspersoners identifikationer af samme VOT-kontinuum, præsenteret dikotisk og binauralt, finder vi samme diffuse fordeling, som vi fandt hos danske forsøgspersoner, jf. figur 3.4.2.

Figur 3.4.2 illustrerer 16 finners identifikationer af de 19 syntetiske vokallyde og er således baseret på 304 præsentationer og identifikationer (16 · 19). Også hos de finske forsøgspersoner finder vi en diffus fordeling af identifikationer: tilsyneladende fremkalder dikotisk præsentation af nogle vokallyde en HØ-præference, mens andre udløser en VØ-præference.

Hvis vi udelukkende fokuserer på forskellen mellem dikotisk præsenterede stimuli og binauralt præsenterede stimuli, er det muligt at isolere variationen i kategoritilskrivningen. Søjlerne i figur 3.4.3 repræsenterer den akkumulerede forskel (i procent) mellem identifikationer ved binaural præsentation og ved dikotisk præsentation med fokuslyden i HØ, og den

akkumulerede forskel mellem identifikationer ved binaural præsentation og ved dikotisk præsentation med fokuslyden i VØ.



Figur 3.4.3

Ved at fremhæve den akkumulerede forskel mellem identifikationer ved binaural præsentation og dikotisk præsentation, kan vi illustrere to væsentlige forhold:

1) For det første afspejler forsøget med dikotisk præsentation af isolerede vokallyde hverken en generel HØ-præference eller VØ-præference. Der er simpelthen ikke grundlag for at hævde, at identifikationer ved binauralt præsenterede vokallyde i højere grad stemmer overens med dikotisk præsenterede fokuslyde i enten HØ eller VØ.

2) For det andet kan vi konkludere, at dikotisk præsentation af vokallyde fremkalder en diffus fordeling af identifikationer. Med få undtagelser afviger de danske identifikationer ved dikotisk præsentation mellem 10 og 30% fra

identifikationer ved binaural præsentation. Hos de finske forsøgspersoner er variationen lidt større, især ved stimulus 9 og 10.

Årsagen til den relativ store forskel mellem dikotisk og binaural præsentation af stimulus 9 og 10 hos de finske forsøgspersoner er sandsynligvis, at stimuli, der er svære at identificere, lettere kan maskeres af en distraktionslyd, end stimuli, der generelt er lette at identificere. Vokallydene er generelt sværere at identificere end konsonantlydene [s], [t] og [d] og er derfor også lettere at maskere, hvilket øger sandsynligheden for variation, dvs. både variation opstået ved tilfældighed og betinget af hemisfæreasymmetri. Da stimulus 9 og 10 indgår i et akustisk kontinuum bestående af ekvidifferente stimuli, er der intet, der taler for, at netop de stimuli skulle være genstand for et særligt identifikationsmønster. Derimod adskiller stimulus 9 og 10 sig fra de resterende ved ikke at tilhøre nogen af de sproglige kategorier: hverken stimulus 9 eller 10 er statistisk signifikante ved $H_0(P < 0,05)$ for nogen af kategorierne, jf. figur 3.4.4. Med andre ord kan variationen i identifikationen af stimulus 9 og 10 let være opstået ved en tilfældighed og ikke ved hemisfæreasymmetri.

Population: Danske forsøgspersoners identifikationer
Kategori: I

Stimulus	n (HØ)	%	X	S	Z	P	n (VØ)	%	X	S	Z	P
1	36	**95**	0,947	**0,226**	-4,19	***	37	**97**	0,974	**0,162**	-6,01	***
2	36	**95**	0,947	**0,226**	-4,19	***	36	**95**	0,947	**0,226**	-4,19	***
3	34	**90**	0,895	**0,311**	-2,88	**0,002***	35	**92**	0,921	**0,273**	-3,37	***
4	27	**71**	0,711	**0,460**	-1,55	**0,061**	30	**79**	0,789	**0,413**	-1,91	**0,028***
5	25	**66**	0,685	**0,481**	-1,37	**0,085**	26	**68**	0,684	**0,471**	-1,45	**0,074**
6	22	**58**	0,579	**0,500**	-1,16	**0,123**	20	**53**	0,526	**0,506**	-1,04	**0,149**
7	11	**29**	0,289	**0,460**	-0,63	**0,264**	13	**34**	0,342	**0,481**	-0,71	**0,239**
8	9	**24**	0,237	**0,431**	-0,55	**0,291**	6	**16**	0,158	**0,370**	-0,43	**0,334**
9	1	**3**	0,026	**0,162**	-0,16	**0,436**	0	**0**	0	**0**	0	**1,0**
10	0	**0**	0	**0**	0	**1,0**	0	**0**	0	**0**	0	**1,0**
11	0	**0**	0	**0**	0	**1,0**	0	**0**	0	**0**	0	**1,0**
12	0	**0**	0	**0**	0	**1,0**	0	**0**	0	**0**	0	**1,0**
13	0	**0**	0	**0**	0	**1,0**	0	**0**	0	**0**	0	**1,0**
14	0	**0**	0	**0**	0	**1,0**	0	**0**	0	**0**	0	**1,0**
15	0	**0**	0	**0**	0	**1,0**	0	**0**	0	**0**	0	**1,0**

16	0	0	0	0	0	1,0	0	0	0	0	0	1,0
17	0	0	0	0	0	1,0	0	0	0	0	0	1,0
18	0	0	0	0	0	1,0	0	0	0	0	0	1,0
19	0	0	0	0	0	1,0	0	0	0	0	0	1,0

Kategori: I

Stimulus	n (HØ)	%	X	S	Z	P	n (VØ)	%	X	S	Z	P
1	2	5	0,053	0,226	-0,24	0,405	1	3	0,026	0,162	-0,16	0,436
2	2	5	0,053	0,226	-0,24	0,405	2	5	0,53	0,226	-0,24	0,405
3	4	11	0,105	0,311	-0,34	0,367	3	8	0,079	0,273	-0,29	0,386
4	11	29	0,289	0,460	-0,63	0,264	8	21	0,211	0,413	-0,51	0,395
5	13	34	0,342	0,481	-0,71	0,239	12	32	0,316	0,471	-0,67	0,251
6	16	42	0,421	0,500	-0,84	0,2	18	47	0,474	0,506	-0,94	0,174
7	27	71	0,711	0,460	-1,55	0,061	25	66	0,658	0,481	-1,37	0,085
8	28	74	0,737	0,446	-1,65	0,05*	30	79	0,789	0,413	-1,91	0,028*
9	35	92	0,921	0,273	-3,37	***	35	92	0,921	0,273	-3,37	***
10	30	79	0,789	0,413	-1,91	0,028*	34	90	0,895	0,311	-2,88	0,002*
11	14	37	0,368	0,489	-0,75	0,227	16	42	0,421	0,500	-0,84	0,20
12	9	24	0,237	0,431	-0,55	0,291	6	16	0,158	0,370	-0,43	0,334
13	2	5	0,053	0,226	-0,24	0,405	0	0	0	0	0	1,0
14	0	0	0	0	0	1,0	0	0	0	0	0	1,0
15	0	0	0	0	0	1,0	0	0	0	0	0	1,0
16	0	0	0	0	0	1,0	0	0	0	0	0	1,0
17	0	0	0	0	0	1,0	0	0	0	0	0	1,0
18	0	0	0	0	0	1,0	0	0	0	0	0	1,0
19	0	0	0	0	0	1,0	0	0	0	0	0	1,0

Kategori: Æ

Stimulus	n (HØ)	%	X	S	Z	P	n (VØ)	%	X	S	Z	P
1	0	0	0	0	0	1,0	0	0	0	0	0	1,0
2	0	0	0	0	0	1,0	0	0	0	0	0	1,0
3	0	0	0	0	0	1,0	0	0	0	0	0	1,0
4	0	0	0	0	0	1,0	0	0	0	0	0	1,0
5	0	0	0	0	0	1,0	0	0	0	0	0	1,0
6	0	0	0	0	0	1,0	0	0	0	0	0	1,0
7	0	0	0	0	0	1,0	0	0	0	0	0	1,0
8	1	3	0,026	0,162	-0,16	0,436	2	5	0,053	0,226	-0,24	0,405

BEST COPY AVAILABLE

9	2	5	0.053	0.226	-0.24	0,405	3	8	0.079	0,273	-0.29	0,386
10	8	21	0.211	0,413	-0.81	0,305	4	11	0.105	0,311	-0.34	0,367
11	24	63	0.632	0,489	-1.29	0,098	22	58	0.579	0,500	-1.16	0,123
12	28	74	0.737	0,446	-1.65	0,05*	30	79	0.789	0,413	-1.91	0,028*
13	34	90	0.895	0,311	-2.88	0,002*	35	92	0.921	0,273	-3.37	***
14	33	87	0.868	0,343	-2.53	0,006*	31	82	0.816	0,393	-2.08	0,019*
15	31	82	0.816	0,393	-2.08	0,019*	31	82	0.816	0,393	-2.08	0,019*
16	25	66	0.658	0,481	-1.37	0,085	28	74	0.737	0,446	-1.65	0,05*
17	19	50	0.5	0,507	-0.99	0,161	20	53	0.526	0,506	-1.04	0,419
18	12	32	0.316	0,471	-0.67	0,251	12	32	0.316	0,471	-0.67	0,251
19	4	11	0.105	0,311	-0.34	0,367	8	21	0.211	0,413	-0.51	0,305

Kategori: A

Stimulus	n (HØ)	%	X	S	Z	P	n (VØ)	%	X	S	Z	P
1	0	0	0	0	0	1,0	0	0	0	0	0	1,0
2	0	0	0	0	0	1,0	0	0	0	0	0	1,0
3	0	0	0	0	0	1,0	0	0	0	0	0	1,0
4	0	0	0	0	0	1,0	0	0	0	0	0	1,0
5	0	0	0	0	0	1,0	0	0	0	0	0	1,0
6	0	0	0	0	0	1,0	0	0	0	0	0	1,0
7	0	0	0	0	0	1,0	0	0	0	0	0	1,0
8	0	0	0	0	0	1,0	0	0	0	0	0	1,0
9	0	0	0	0	0	1,0	0	0	0	0	0	1,0
10	0	0	0	0	0	1,0	0	0	0	0	0	1,0
11	0	0	0	0	0	1,0	0	0	0	0	0	1,0
12	1	3	0,026	0,162	-0.16	0,436	2	5	0,053	0,226	-0,24	0,405
13	2	5	0.053	0,226	-0.24	0,405	3	8	0.079	0,273	-0,29	0,386
14	5	13	0.132	0,343	-0.39	0,348	7	18	0.184	0,393	-0,47	0,319
15	7	18	0.184	0,393	-0.47	0,319	7	18	0.184	0,393	-0,47	0,319
16	13	34	0.342	0,481	-0.71	0,239	10	26	0.263	0,446	-0.59	0,278
17	19	50	0.5	0,507	-0,99	0,161	18	47	0.474	0,506	-0.94	0,174
18	26	68	0.684	0,471	-1.45	0,074	26	68	0.684	0,471	-1.45	0,074
19	34	90	0.895	0,311	-2.88	0,002*	30	79	0.789	0,413	-1.91	0,028*

Population: Finske forsøgspersoners identifikationer

Kategori: I

Stimulus	n (HØ)	%	X	S	Z	P	n (VØ)	%	X	S	Z	P
1	16	100	1,0	0	0	•••	16	100	1,0	0	0	•••
2	16	100	1,0	0	0	•••	16	100	1,0	0	0	•••
3	16	100	1,0	0	0	•••	16	100	1,0	0	0	•••
4	16	100	1,0	0	0	•••	16	100	1,0	0	0	•••
5	15	94	0,938	0,25	-3,75	•••	16	100	1,0	0	0	•••
6	15	94	0,938	0,25	-3,75	•••	13	81	0,813	0,403	-2,02	0,022*
7	15	94	0,938	0,25	-3,75	•••	12	75	0,75	0,447	-1,68	0,046*
8	12	75	0,75	0,447	-1,68	0,046*	11	69	0,688	0,479	-1,44	0,075
9	9	56	0,563	0,512	-1,45	0,074	7	44	0,438	0,512	-1,13	0,129
10	5	31	0,313	0,479	-0,65	0,258	5	31	0,313	0,479	-0,65	0,258
11	0	0	0	0	0	1,0	2	13	0,125	0,342	-0,37	0,356
12	0	0	0	0	0	1,0	0	0	0	0	0	1,0
13	0	0	0	0	0	1,0	0	0	0	0	0	1,0
14	0	0	0	0	0	1,0	0	0	0	0	0	1,0
15	0	0	0	0	0	1,0	0	0	0	0	0	1,0
16	0	0	0	0	0	1,0	0	0	0	0	0	1,0
17	0	0	0	0	0	1,0	0	0	0	0	0	1,0
18	0	0	0	0	0	1,0	0	0	0	0	0	1,0
19	0	0	0	0	0	1,0	0	0	0	0	0	1,0

Kategori: E

Stimulus	n (HØ)	%	X	S	Z	P	n (VØ)	%	X	S	Z	P
1	0	0	0	0	0	1,0	0	0	0	0	0	1,0
2	0	0	0	0	0	1,0	0	0	0	0	0	1,0
3	0	0	0	0	0	1,0	0	0	0	0	0	1,0
4	0	0	0	0	0	1,0	0	0	0	0	0	1,0
5	1	6	0,063	0,25	-0,25	0,401	0	0	0	0	0	1,0
6	1	6	0,063	0,25	-0,25	0,401	3	19	0,188	0,403	-0,47	0,319
7	1	6	0,063	0,25	-0,25	0,401	4	25	0,25	0,447	-0,56	0,288
8	4	25	0,25	0,447	-0,56	0,288	5	31	0,313	0,479	-0,65	0,258
9	7	44	0,438	0,512	-1,13	0,129	9	56	0,563	0,512	-1,45	0,074
10	11	69	0,688	0,479	-1,44	0,075	11	69	0,688	0,479	-1,44	0,075
11	15	94	0,938	0,25	-3,75	•••	14	88	0,875	0,342	-2,56	0,005*
12	15	94	0,938	0,25	-3,75	•••	16	100	1,0	0	0	•••

BEST COPY AVAILABLE

13	15	94	0,938	0,25	-3,75	***	16	100	1,0	0	0	***
14	12	75	0,75	0,447	-1,68	0,046*	14	88	0,875	0,342	-2,56	0,005*
15	12	75	0,75	0,447	-1,68	0,046*	11	69	0,688	0,479	-1,44	0,075
16	8	50	0,5	0,516	-0,97	0,66	10	63	0,625	0,50	-1,25	0,106
17	7	44	0,438	0,512	-1,13	0,129	8	50	0,50	0,516	-0,97	0,166
18	5	31	0,313	0,479	-0,65	0,258	5	31	0,313	0,479	-0,65	0,258
19	2	13	0,125	0,342	-0,37	0,356	3	19	0,188	0,404	-0,47	0,319

Kategori: A

Stimulus	n (HØ)	%	X	S	Z	P	n (VØ)	%	X	S	Z	P
1	0	0	0	0	0	1,0	0	0	0	0	0	1,0
2	0	0	0	0	0	1,0	0	0	0	0	0	1,0
3	0	0	0	0	0	1,0	0	0	0	0	0	1,0
4	0	0	0	0	0	1,0	0	0	0	0	0	1,0
5	0	0	0	0	0	1,0	0	0	0	0	0	1,0
6	0	0	0	0	0	1,0	0	0	0	0	0	1,0
7	0	0	0	0	0	1,0	0	0	0	0	0	1,0
8	0	0	0	0	0	1,0	0	0	0	0	0	1,0
9	0	0	0	0	0	1,0	0	0	0	0	0	1,0
10	0	0	0	0	0	1,0	0	0	0	0	0	1,0
11	1	6	0,063	0,25	-0,25	0,401	0	0	0	0	0	1,0
12	1	6	0,063	0,25	-0,25	0,401	0	0	0	0*	0	1,0
13	1	6	0,063	0,25	-0,25	0,401	0	0	0	0	0	1,0
14	4	25	0,25	0,447	-0,56	0,288	2	13	0,125	0,342	-0,37	0,356
15	4	25	0,25	0,447	-0,56	0,288	5	31	0,313	0,479	-0,65	0,258
16	8	50	0,5	0,516	-0,97	0,166	6	38	0,375	0,5	-0,75	0,227
17	9	56	0,563	0,512	-1,45	0,074	8	50	0,5	0,516	-0,97	0,166
18	11	69	0,688	0,479	-1,44	0,075	11	69	0,688	0,479	-1,44	0,075
19	14	88	0,875	0,342	-2,56	0,005*	13	81	0,813	0,403	-2,02	0,022*

n: Observationsværdi

%: Procent

X: Den aritmetiske middelværdi

S: Standardafvigelsen

Z: Observationsværdiens afvigelse fra den aritmetiske middelværdi udtrykt i standardafvigelsesenheder

P:	Sandsynlighed for H_0
*:	$P < 0.05$
***:	$P < 0.001$

Figur 3.4.4

3.5 Resultater fra forsøg med dikotisk lytning: [la], [ma], [na], [fa] og [va]

I de to foregående kapitler så vi, at perceptionen af dikotisk præsenterede stimuli fra et VOT-kontinuum afspejler en HØ-præference, hvorimod stimuli fra et formantkontinuum ikke fremkalder nogen ørepræference, dvs. hverken en HØ- eller en VØ-præference. Jeg har valgt at tolke disse resultater som et udtryk for hemisfæreasymmetri hos både de danske og de finske forsøgspersoner.

For endvidere at underbygge hypotesen om specifikke lateraliserede sprogprocesser, der varetager analysen af artikulatorisk komplekse sproglyde, har jeg udvidet undersøgelsen til også at omfatte CV-stavelser bestående af en lateral, nasaler og frikativer kombineret med et langt [a]. Formålet med at præsentere disse stimuli er naturligvis at finde ud af, om lateraler, nasaler og frikativer resulterer i en dikotisk effekt, dvs. en HØ- eller VØ-præference. Hvis analysen af likvider, frikativer og semivokaler helt eller delvis afhænger af de lateraliserede processer, kan vi nemlig forvente af finde en HØ-præference i identifikationen af de dikotisk præsenterede CV-stavelser.

Forsøget med dikotisk præsentation af [la], [ma], [na], [fa] og [va] blev kun udført med 37 forsøgspersoner, der alle er danskere. Hver stimulus bestod af en fokuslyd, én af CV-stavelserne, og en distraktionslyd. De 37 forsøgspersoner blev bedt om at identificere hver fokuslyd to gange, hvilket betyder, at det empiriske materiale består af 370 præsentationer og identifikationer (37 · 5 · 2).

Figur 3.5.1

156

Figur 3.5.1 illustrerer de 37 danske forsøgspersoners identifikationer af de 5 CV-stavelser.

Som det fremgår af figuren har forsøgspersonerne i et vist omfang været i stand til at identificere stimuli som tilhørende én af de fem kategorier. De har dog haft relativt lettere ved identificere [la], [ma] og [fa] end [na] og [va].

Det fremgår desuden, at forsøgspersonerne oftere har kunnet identificere CV-stavelserne, når de er blevet præsenteret som fokuslyd for HØ end for VØ. I et enkelt tilfælde, nemlig ved præsentationen af [fa], har forsøgs-personerne identificeret stavelsen lige ofte i HØ og VØ. Generelt kan vi dog konkludere, at også ved præsentation af [la], [ma], [na], [fa] og [va] afslører forsøg med DL en om end beskeden HØ-præference.

Sådan som jeg tidligere opstillede hypotesen om en lateraliseret sprog-processor (se side 119f), var forudsigelsen, at en dikotisk præsentation af lateraler, nasaler og frikativer ville resultere i en dikotisk effekt, men en mindre HØ-præference end ved dikotisk præsentation af klusilerne [s], [t] og [d]. På grund af forskelle mellem de to forsøgsopstillinger har jeg desværre ikke mulighed for at foretage en direkte kvantitativ sammenligning af den dikotiske effekt ved præsentationen af de to typer stimuli. Derimod er der mulighed for at sammenligne forskellen mellem identifikationer ved præsentation for hhv. HØ og VØ. I en beregning af den kumulerede forskel (i procent) mellem korrekte identifikationer ved præsentation for HØ og VØ viser det sig, at hver stimulus fra det akustiske VOT-kontinuum i gennemsnit er blevet korrekt identificeret 7,8% oftere ved præsentation for HØ end for VØ[1]. Omvendt udgør den kumulerede forskel mellem korrekte identifikationer af [la], [ma], [na], [fa] og [va] ved præsentation for HØ og VØ kun 3,6%, hvilket vil sige, at disse stimuli i gennemsnit er blevet korrekt identificeret 3,6% oftere ved præsentation for HØ end for VØ. Med andre ord kan jeg altså konkludere, at dette forsøg med dikotisk præsentation af [la], [ma], [na], [fa] og [va] afslører en generel HØ-præference, men en mindre HØ-præference end i et tilsvarende forsøg med et akustisk VOT-kontinuum.

[1] For overhovedet at kunne foretage denne sammenligning har jeg været nødt til at definere den "korrekte" kategoritilskrivning for hver stimulus, hvilket jeg har gjort på baggrund af den højeste procentvise kategoritilskrivning i identifikationsforsøgene. Beregningen af den kumulerede forskel er iøvrigt udelukkende foretaget på baggrund af data fra gruppen af danske forsøgspersoner.

Population: danske forsøgspersoners identifikationer

Stimulus	n (HØ)	%	X	S	Z	P	n (VØ)	%	X	S	Z	P
[la]	74	100	1,0	0	0	***	71	96	0,959	0,20	-1,8	***
[ma]	72	97	0,973	0,163	-5,97	***	67	91	0,905	0,295	-3,07	***
[na]	39	53	0,527	0,503	-1,05	0,147	37	50	0,50	0,503	-0,99	0,161
[fa]	67	91	0,905	0,295	-3,07	***	67	91	0,905	0,295	-3,07	***
[va]	20	27	0,270	0,447	-0,60	0,274	16	22	0,216	0,414	-0,52	0,302

n: Observationsværdi

%: Procent

X: Den aritmetiske middelværdi

S: Standardafvigelsen

Z: Observationsværdiens afvigelse fra den aritmetiske middelværdi udtrykt i standardafvigelsesenheder

P: Sandsynlighed for H_0

*: $P < 0,05$

***: $P < 0,001$

Figur 3.5.2

Generelt resulterer forsøg med dikotisk præsentation af lateraler, nasaler og frikativer i en HØ-præference – dog en mindre HØ-præference end ved præsentation af klusiler og en større HØ-præference end ved præsentation af vokaler. Resultatet fra dette forsøg afviger ikke fra denne tendens.

Hidtil har jeg forklaret den kategoriske og dikotiske effekt ved præsentation af klusiler og vokaler ud fra dikotomien artikulatorisk kontinuitet/diskontinuitet. Spørgsmålet må nu være, hvordan [l], [m], [n], [f] og [v] relateres til disse termer, så de stadig adskiller sig fra både vokaler og klusiler. Traditionelle artikulatoriske beskrivelser, der tager udgangspunkt i måde/åbningsgrad, angiver ingen tilstrækkelige artikulatoriske karakteristika for ovenstående perceptuelle kategori: på den ene side kategoriseres frikativer og lukkelyde som obstruenter; på den anden side kategoriseres nasaler og lateral(-er) som sonoranter (jf. f.eks. Heger 1992). Derimod kan den perceptuelle kategori beskrives som en medierende kategori, der besidder såvel vokalagtige som konsonantagtige (artikulatoriske) egenskaber: [l], [m], [n], [f] og [v] artikuleres alle med en vis hindring for udåndingsluften langs mundkanalens midterlinje, hvilket er et træk de har tilfælles med klusilerne;

153

omvendt er de ikke karakteriseret ved abrupt lukke eller åbning under selve artikulationen, hvilket er et træk de har tilfælles med kontekstfrie vokaler.

4 Konklusion

Gennem de sidste 45 år har en lang række undersøgelser med kategorisk perception (KP) vist, at visse sproglyde perciperes som diskrete størrelser og derved fremkalder en såkaldt 'phoneme boundary effect'. Andre sproglyde perciperes mere kontinuerligt og synes ikke i samme grad at fremkalde skarpe kategorigrænser. Resultaterne af undersøgelserne viser, at lukkelyde fremkalder den kraftigste kategoriske effekt, nasaler, lateraler og frikativer en mindre kategorisk effekt, og at vokaler perciperes kontinuerligt.

Undersøgelser fra et helt andet paradigme, nemlig dikotisk lytning, har ligeledes gennem de sidste 45 år vist, at dikotisk præsentation af visse sproglyde kan fremkalde højreøre-præference. Denne HØ-præference tolkes almindeligvis som resultat af en hemisfæreasymmetri. Det er dog ikke alle sproglyde, som kan fremkalde en HØ-præference. Resultaterne af undersøgelserne viser, at lukkelyde fremkalder den største HØ-præference, likvider, frikativer og semivokaler en mindre HØ-præference og at vokaler ikke fremkalder en ørepræference.

I begge forskningsparadigmer arbejdes der ud fra en hypotese om, at perceptionen af sprog involverer et særligt sprogmodul, en såkaldt sprog-processor, der varetager analysen af artikulatorisk komplekse sproglyde. I den ene forskningstradition (KP) underbygges hypotesen af forsøg, der viser, at klusiler perciperes som diskrete størrelser og vokaler mere kontinuerligt. I den anden forskningstradition anses hypotesen for at være den eneste plausible forklaring på, at forsøg med DL afslører en hemisfæreasymmetri ved præsentation af klusiler, men ikke ved vokallyde. I begge forskningsparadigmer har man dog hidtil kun gjort sig teoretiske overvejelser om en eventuel sammenhæng mellem den kategoriske og den dikotiske effekt.

Med nærværende undersøgelse har jeg forsøgt at kaste lys på ovenstående problemstilling. Indledningsvis producerede jeg et VOT-kontinuum og et formantkontinuum. Disse kontinua skulle udgøre hhv. konsonant- og vokal-stimuli i to forskellige forsøg, nemlig et med KP og et andet med DL.

I forsøget med KP afslørede det artikulatorisk diskontinuerlige konsonant-kontinuum en høj grad af kategorisk effekt: i identifikationsforsøget inddelte forsøgspersonerne det akustiske kontinuum i skarpt afgrænsede kategorier; i

diskriminationsforsøget kunne forsøgspersonerne ikke skelne mellem stimuli, som de havde tilskrevet samme kategori, hvorimod de meget lettere kunne skelne mellem stimuli fra forskellige kategorier. Da samme akustiske kontinuum blev præsenteret i forsøget med DL, viste forsøgspersonerne en markant bedre evne til at identificere stimuli præsenteret for HØ end for VØ, hvilket kan tolkes som et resultat af hemisfæreasymmetri.

Omvendt fremkaldte det syntetiske vokalkontinuum slet ingen kategorisk effekt: Forsøgspersonerne var i stand til relativt konsekvent at tilskrive stimuli de samme kategorier, men i diskriminationsøvelsen gjorde det ingen forskel for forsøgspersonerne, om de skulle diskriminere stimuli fra samme eller forskellige kategorier – resultatet var lige ringe. Ligeledes fandt jeg heller ikke nogen generel dikotisk effekt, da jeg præsenterede det samme vokal-kontinuum i forsøget med DL: Forsøgspersonerne var hverken generelt bedre til at identificere stimuli præsenteret for HØ end VØ eller omvendt.

Da flere forsøg har vist en tendens til en mindre kategorisk og dikotisk effekt ved præsentation af især likvider og frikativer, besluttede jeg også at inddrage CV-stavelserne [la], [ma], [na], [fa] og [va]. Disse stimuli blev præsenteret i forsøget med DL og viste – som forventet – en HØ-præference, dog mindre end den jeg fandt ved præsentation af det akustiske VOT-kontinuum.

På baggrund af nærværende undersøgelse kan jeg konkludere, at der eksisterer en klar sammenhæng mellem graden af kategorisk effekt og effekten i forsøg med DL, dvs. hemisfæreasymmetri.

For endvidere at fastslå om egenskaberne 'kontinuitet' og 'diskontinuitet' vedrører sproglydes akustiske struktur eller om de snarere knytter sig til den sprogspecifikke artikulatoriske struktur, valgte jeg også at inddrage en gruppe finske forsøgspersoner i undersøgelsen. Formålet med den komparative metode var at finde ud af, om forsøgspersoner, der taler forskellige modersmål, kategoriserer de samme akustiske kontinua ens eller forskelligt. Hvis den kategoriske effekt skyldes sproglydenes akustiske struktur, skulle man forvente, at den determinerer de fonetiske kategorier – dvs. identifikations- og diskriminationsgrænserne – uafhængigt af den lingvistiske organisering af forsøgspersonens modersmål. Hvis den kategoriske effekt derimod skyldes sprogspecifikke artikulatoriske strukturer, må den forventes at forkomme ved alle lingvistiske kategorigrænser – og kun dér.

Spørgsmålet om hvorvidt den kategoriske effekt skyldes psykoakustiske faktorer eller sprogspecifikke fonetiske kategoriseringer kan ikke besvares helt éntydigt på baggrund af forsøgsresultaterne. Det fremgår dog meget tydeligt, at den lingvistiske organisering af forsøgspersonernes modersmål har stor indflydelse på identifikationen og diskriminationen af sproglyde fra akustiske kontinua; de to grupper af forsøgspersoner strukturerer de akustiske kontinua forskelligt, men helt i overensstemmelse med de fonetiske kontraster på de respektive sprog. Omvendt viser resultater fra de danske forsøgspersoner, at sprogspecifikke artikulatoriske strukturer ikke kan forklare alle fænomener; til trods for at stemthedsgrænsen (0 ms VOT) ikke indgår i den danske strukturering af udtrykssubstansen, afslører forsøg med danske forsøgspersoner en øget diskrimination af stimuli fra samme kategori, men fra hver sin side af stemthedsgrænsen. Meget tyder altså på, at egenskaberne 'kontinuitet' og 'diskontinuitet' bedst kan appliceres på de sprogspecifikke artikulatoriske strukturer, men at sprogbrugere også anvender anden information end de artikulatoriske kommandoer i identifikationen af sproglyde – uanset om de er karakteriseret ved artikulatorisk kontinuitet eller diskontinuitet.

Litteraturliste.

Abramson, A.S. 1961. Identification and discrimination of phonemic tones. JASA 33: 842.

Abramson, A.S. & Lisker, L. 1970. Discriminability along the voicing continuum: Cross-language tests. Proceedings of the sixth international congress of phonetic sciences. Prague: Academia, 569-573.

Ades, A.E. 1974. How phonetic is selective adaptation? Experiments on syllabic position and vowel environment. P&P 16: 61-66.

Ades, A.E. 1976. Adapting the property detectors for speech perception. I Wales, R.J. & Walker, E. (ed.): New approaches to language mechanisms. Amsterdam: North-Holland.

Ades, A.E. 1977. Vowels, consonants, speech, and nonspeech. PR 84: 524-530.

Bally, C. & Sechehaye, A. (ed.). 1959. Course in General Linguistics, Ferdinand de Saussure. Translated by Wade Baskin. McGraw

Basbøll, H. 1977. Dansk fonetik og fonologi. Skitse til systematisk indføring. Odense Universitet.

Basbøll, H. & Wagner, J. 1985. Kontrastive Phonologie des Deutschen und Dänishen Segmentale Wortphonologie und -phonetik. Lingvistische Arbeiten 160. Tübingen: Max Niemeyer Verlag.

Basbøll, H. 1997. On Danish Schwa. I Bertinetto, P.M. et al (ed.): Certamen phonologicum III. Torino: Rosenberg & Sellier, 3-26.

Basbøll, H. 1999. Syllables in Danish. I Hulst, H.v d. & Ritter, N. (ed.): Syllables: Views and Facts. Mouton de Gruyter.

Bastian, J. & Abramson, A.S. 1962. Identification and discrimination of phonemic vowel duration. JASA 34: 743.

Bastian, J. & Abramson, A.S. 1964. Identification and discrimination of phonemic vowel duration. Speech research and instrumentation. NY: Haskins Laboratories.

Bastian, J., Eimas, P.D. & Liberman, A.M. 1961. Identification and discrimination of a phonemic contrast induced by silent interval. JASA 33: 842.

Bates, E. (in press). On the Nature and Nurture of Language. Bizzi, E., Calissano, P. & Volterra, V. (ed.): Frontiere della biologia. The brain of homo sapiens. Rome: Giovanni Trecanni.

Berlin, B. & Kay, P. 1969. Basic color terms: Their universality and evolution. University of California Press

Berlin, C.I., Hughes, L.F., Lowe-Bell, S.S., Berlin, H.L. 1973. Dichotic right ear advantage in children 5 to 13. Cortex 9: 393-402.

Blumstein, S.E. & Cooper, W.E. 1974. Hemispheric processing of intonation contours. Cortex 10: 146-158.

Bond, Z.S. 1981. From an Acoustic Stream to a Phonological Representation: The Perception of Fluent Speech. In Lass, N.J. (ed.) Speech and Language: Advances in basic research and practice 6: 375-410. NY: Academic.

Borden, G. J. & Harris, K. 1984. Speech Science Primer. Baltimore, MD: Williams and Wilkins.

Bottini, G., Corcoran, R., Sterzi, R., Paulesu, E., Schenone, P., Scarpa, P., Frackowiak, R.S.J. & Frith, C.D. 1994. The role of the right hemisphere in the interpretation of figurative aspect of language. A positron emission tomography activation study. Brain 117: 1241-1253.

Brink, L. et al. 1991. Den store danske udtaleordbog. Kbh.: Munksgaard 1991.

Broadbent, D.E. 1954. The role of auditory localization in attention and memory span. JEP 47(3): 191-196.

Bryden, M.P. & Ley, R.G. 1983. Right hemispheric involvement in imagery and affect. I Perecman, E. (ed.): Cognitive processing in the right hemisphere. Orlando, Florida: Academic Press, 111-123.

Chomsky, N. 1965. Aspects of the Theory of Syntax. Cambridge, Mass.: MIT Press.

Chomsky, N. & Halle, M. 1968. The Sound Pattern of English. N.Y.: Harper & Row.

Clark, H.H. & Clark, E.V. 1977. Psychology and Language. An introduction to psycholinguistics. New York: Harcourt Brace Jovanovich.

Cole, R.A. 1977. Invariant features and feature detectors: Some developmental implications. I Segalowitz, S.J. & Gruber, F.A. (ed.): Language Development and Neurological Theory.

Cooper, F.S., Delattre, P.C., Liberman, A.M., Borst, J.M. & Gerstman, L.J. 1952. Some experiments on the perception of synthetic speech sounds. JASA 24: 597-606.

Cooper, W.E. 1974a. Contingent feature analysis in speech perception. P&P 16: 201-204.

Cooper, W.E. 1974b. Adaptation of phonetic feature analyzers for place of articulation. JASA 56: 617-627.

Cooper, W.E. 1974c. Selective adaptation for acoustic cues of voicing in initial stops. Journal of phonetics 2: 303-313.

Crystal, D. 1995. A dictionary of linguistics and phonetics. Oxford: Blackwell Publishers.

Crystal, I. & House, A.S. 1974. The effect of local signal level on differential performances in dichotic listening. JASA 55: 434.

Curry, F.K.W. 1967. A comparison of left-handed and right-handed subjects on verbal and non-verbal dichotic listening tasks. Cortex 3: 343-352.

Curry, F.K.W. 1968. A comparison of the performances of a right hemispherectomized subject and 25 normals on four dichotic listening tasks. Cortex 4: 144-153.

Cutting, J.E. 1972. Ear advantage for stops and liquids in initial and final position. Haskins Lab. Status Report SR-31/32: 57-66.

BEST COPY AVAILABLE

Cutting, J.E. 1974. Two left-hemisphere mechanisms in speech perception. P&P 16(3): 601-612.
Cutting, J.E. & Rosner, B.S. 1974. Categories and boundaries in speech and music. P&P 16: 564-570.
Cutting, J.E., Rosner, B.S. & Foard, C.F. 1976. Perceptual categories for musiclike sounds: Implications for theories of speech perception. JEP 28: 361-378.
Darwin, C.J. 1971. Ear differences in the recall of fricatives and vowels. JEP 23: 46-62.
Day, R. & Vigorito, J.M. 1973. A parrallel between encodedness and the ear advantage: evidence from a temporal-order judgement task. JASA 53: 368.
Delattre, P.C., Liberman, A.M. & Cooper, F.S. 1955. Acoustic Loci and Transitional Cues for Consonants. JASA 27: 769-773.
Dirks, D. 1964. Perception of dichotic and monaural verbal material and cerebral dominance for speech. Acta Otolaryng 58: 73-80.
Dorman, M.F., Raphael, L.J., Liberman, A.M. 1979. Some experiments on the sound of silence in phonetic perception. JASA 65(6): 1518-1532.
Eimas, P., Cooper, W.E. & Corbit, J.D. 1973. Some properties of linguistic feature detectors. P&P 13: 247-252.
Eimas, P. & Corbit, J.D. 1973. Selective adaptation of linguistic feature detectors. Cognitive Psychology 4: 99-109.
Eimas, P.D. 1963. The relation between identification and discrimination along speech and non-speech continua. Language and Speech 6: 206-217.
Eimas, P.D. 1974a. Linguistic processing of speech by young infants. I Schiefelbusch, R. & Lloyd, L. (ed.): Language perspectives - acquisition, retardation, and intervention. Baltimore: University Park Press.
Eimas, P.D. 1974b. Auditory and linguistic processing of cues for place of articulation by infants. P&P 16: 513-521.
Eimas, P.D. 1975a. Auditory and phonetic coding of the cues for speech: Discrimination of the r-l distinction by young infants. P&P 18: 341-347.
Eimas, P.D. 1975b. Developmental studies of speech perception. I Cohen, L.B. & Salapatek (ed.): Infant perception. NY:Academic Press.
Eimas, P.D., Siqueland, E.R., Jusczyk, P. & Vigorito, J. 1971. Speech perception in early infancy. Science 171: 304-306.
Fant, G. 1958. Modern instruments and methods for acoustic studies of speech. Proceedings of the 8th international congress of linguistics. Oslo.
Fischer-Jørgensen, E. 1954. Acoustic analysis of stop consonants. Miscellanea Phonetica 2: 42-59.
Fischer-Jørgensen, E. 1962. Almen fonetik. Kbh.: Rosenkilde og Bagger.
Fischer-Jørgensen, E. 1966. Phonetic analysis of Danish stop consonants. ARIPUC 1: 31-34.
Fischer-Jørgensen, E. 1981. Fonetik, studiet af sproglyde. Kbh.: Det Kongelige Danske Videnskabernes Selskab, Grundvidenskaben i dag, vol. 27.
Fitch, H.L. et al. 1980. Perceptual equivalence of two acoustic cues for stop-consonant manner. P&P 27: 343-350.
Foss, D.J. & Hakes, D.T. 1978. Psycholinguistics: an introduction to the psychology of language. Englewood Cliffs, N.J.: Prentice-Hall.
Fowler, C.A. 1986. An event approach to the study of speech perception from a direct-realist perspective. Journal of Phonetics 14: 3-28.
Frandsen, N. 1994a. Empiriske undersøgelser af hemisfæreforskelle i forbindelse med kognitiv lingvistik: Hvorfor og Hvordan. Status over et års forsøg. Ikke publiceret.
Frandsen, N. 1994b. Freuds model af psyken. Menneske & Natur. Arbejdspapir 51.
Frøkjær-Jensen, B. 1966. The Danish long vowels. ARIPUC 1: 34-47.
Fry, D.B., Abramson, A.S., Eimas, P.D. & Liberman, A.M. 1962. The identification and discrimination of synthetic vowels. Language and Speech 5: 171-189.
Fujisaki, H. & Kawashima, T. 1969. On the modes and mechanisms of speech perception. Annual report of the Engineering research institute, Faculty of engineering, University of Tokyo, 28: 67-73.
Fujisaki, H. & Kawashima, T. 1970. Some experiments on speech perception and a model for the perceptual mechanism. Annual report of the Engineering research institute, Faculty of engineering, University of Tokyo, 29: 207-214.
Fujisaki, H. & Kawashima, T. 1971. A model of the mechanisms for speech perception: Quantitative analysis of categorical effects in discrimination. Annual report of the Engineering research institute, Faculty of engineering, University of Tokyo, 30: 59-68.
Garman, M. 1990. Psycholinguistics. Cambridge University Press.
Gibson, J.J. 1966. The senses considered as perceptual systems. Mass.: Houghton-Mifflin.
Haggard, M. 1971. Encoding and the REA for speech signals. JEP 23: 34-45.
Halle, M. 1964. On the basis of phonology. I Fodor, J.A. & Katz, J.J. (ed.): The structure of language. Englewood Cliffs, N.J.: Prentice-Hall.
Halle, M. & Stevens, K. N. 1959. Analysis by synthesis. I Wathen-Dunn, W. & Woods, L.E. (ed.): Proceedings of the seminar on speech comprehension and processing 2: D7.

Halle, M. & Stevens, K. N. 1972. On phonetic features. Conference on speech analysis and synthesis, 194-197.
Hansen, P.M. 1990. Udtaleordbog/Dansk udtale. Kbh.: Gyldendal.
Harley, T.A. 1995. The Psychology of Language. From Data to Theory. Erlbaum (UK) Taylor & Francis. Psychology Press.
Harris, K.S., Hoffman, H.S., Liberman, A.M., Delattre, P.C. & Cooper, F.S. 1958. Effect of third-formant transitions on the perception of the voiced stop consonants. JASA 30: 122-126.
Hartshorne, C. & Weiss, P. (ed.) 1935. The Collected Papers of Charles Sanders Peirce 3. Cambridge MA: Harvard University Press.
Hayden, M.E., Kirstein, E. & Singh, S. 1979. Role of distinctive features in dichotic perception of 21 English consonants. JASA 65(4): 1039-1047.
Heger, S. 1992 (1974, 1975, 1981, 1992). Sprog & lyd. Kbh.: Akademisk Forlag.
Hjelmslev, L. 1943. Omkring Sprogteoriens Grundlæggelse. Kbh.
Hjelmslev, L. 1954. Almindelig fonetik. I Andersen, P. & Hjelmslev, L: Fonetik. 233-308. Kbh.: Rosenkilde & Bagger.
Jakobson, R., Fant, G. & Halle, M. 1963. Preliminaries to speech analysis: The distinctive features and their correlates. Cambridge, Mass.: MIT Press. Også publiseret under samme titel i Technical Report 13, Acoustic Laboratories, Cambridge, Mass.: MIT Press, 1952.
Jakobson, R. & Halle, M. 1956. Fundamentals of language. The Hague: Mouton.
Jensen, J.N. & Madsen, T.K. 1995. En undersøgelse af dikotisk lytning som metode til studiet af hemisfæreasymmetri og sprogperception. Eksamensopgave i sproglig metode og analyse. Ikke publiceret.
Johnson, M. 1987. The body in the mind. The bodily basis of meaning, imagination, and reason. The University of Chicago Press.
Keller, E. 1994. Signalyze. Signal analysis for speech and sound. Ma.: Network Technology Corporation.
Kimura, D. 1961a. Cerebral dominance and the perception of verbal stimuli. Canadian Journal of Psychology 15: 166-171.
Kimura, D. 1961b. Some effects of temoral-lope damage on auditory perception. Canadian Journal of Psychology 15: 156-165.
Kimura, D. 1964. Left-right differences in the perception of melodies. JEP 16: 355-358.
Kimura, D. 1967. Functional asymmetry of the brain in dichotic listening. Cortex 3: 163-178.
Klatt, D.H. 1979. Speech perception: A model of acoustic-phonetic analysis and lexical access. Journal of Phonetics 7: 279-312.
Klatt, D.H. 1989. Lexical representations for speech production and perception. I Myers, T., Laver, J. & Anderson, J. (ed.): The cognitive representation of speech. 11-31. Amsterdam: North-Holland.
Krashen, S.D. 1977. Venstre hemisfære. I Wittrock M.C. (ed.): Menneskets dobbelthjerne. En artikelsamling. 111-134. NNF.
Kuhl, P. 1976. Speech perception by the chinchilla: Categorical perception of synthetic alveolar plosive consonants. 92nd meeting of the Acoustical Society of America. San Diego.
Kuhl, P. & Miller, J.D. 1975. Speech perception by the chinchilla: Voiced-voiceless distinction in alveolar plosive consonants. Science 190: 69-72.
Kuhl, P. & Miller, J.D. 1978. Speech perception by the chinchilla: Identification functions for synthetic VOT stimuli. JASA 63: 905-917.
Ladefoged, P. & Broadbent, D.E. 1957. Information conveyed by vowels. JASA 29: 98-104.
Lakoff, G. 1987. Women, Fire, and Dangerous Things. What categories reveal about the mind. Chicago.
Lane, H. 1965. Motor theory of speech perception: A critical review. PR 72: 275-309.
Larkey, L.S., Wald, J. & Strange, W. 1978. Perception of synthetic nasal consonants in initial and final syllable position. P&P 23: 299-312.
Lasky, R., Syrdal-Lasky, A. & Klein, D. 1975. VOT discrimination by four to six month old infants from Spanish environments. Journal of Experimental Child Psychology 20: 215-225.
Lass, R. 1984. Phonology: an introduction to basic concepts. Cambridge University Press.
Lehiste, I. 1967. Readings in Acoustic Phonetics. Cambridge, MA: MIT Press.
Lehiste, I. & Peterson, G.E. 1959. Vowel amplitude and phonemic stress in american english. JASA 31: 428-435.
Lenneberg, E.H. 1962. Understanding language without ability to speak: A case report. Journal of Abnormal and Social Psychology 65: 419-425.
Liberman, A.M. 1970. The grammars of speech and language. Cognitive Psychology 1: 301-323.
Liberman, A.M., Cooper, F.S., Shankweiler, D., Studdert-Kennedy, M. 1967. Perception of the speech code. PR 74(6): 431-461.
Liberman, A.M., Delattre, P.C. & Cooper, F.S. 1952. The Role of Selected Stimulus-Variables in the Perception of the Unvoiced Stop Consonants. American Journal of Psychology 65: 497-516.
Liberman, A.M., Delattre, P.C., Cooper, F.S. & Gerstman, L.J. 1954. The Role of Consonant-Vowel Transitions in the Perception of the Stop and Nasal Consonants. Psychological Monographs: General

and Applied 68: 1-13.
Liberman, A.M., Harris, K.S., Eimas, P.D., Lisker, L. & Bastian, J. 1961. An effect of learning on speech perception: The discrimination of duration of silence with and without phonemic significanse. Language and Speech 4: 175-195.
Liberman, A.M., Harris, K.S., Hoffman, H.S. & Griffith, B.C. 1957. The discrimination of speech sounds within and across phoneme boundaries. JEP 54(5): 358-368.
Liberman, A.M., Harris, K.S., Kinney, J.A. & Lane, H. 1961. The discrimination of relative onset-time of the components of certain speech and non-speech patterns. JEP 61: 379.
Liberman, A.M., Ingeman, F., Lisker, L., Delattre, P.C. & Cooper, F.S. 1959. Minimal rules for synthesizing speech. JASA 31: 1490-1499.
Liberman, A.M. & Mattingly, I.G. 1985. The motor theory of speech perception revised. Cognition 21: 1-36.
Lieberman, P. 1974. On the evolution of language: A unified view. Cognition 2(1): 59-94.
Lieberman, P. & Blumstein, S.E. 1988. Speech physiology, speech perception, and acoustic phonetics. Cambridge, NY.
Lisker, L. 1957. Closure duration and the intervocalic voiced-voiceless distinction in English. Language 33: 42-49.
Lisker, L & Abramson, A.S. 1970. The voicing dimension: Some experiments in comparative phonetics. Proceedings of the sixth international congress of phonetic sciences. 563-567. Prague: Academia.
Luria, A.R. 1975. Hjernen - En introduktion til neuropsykologien. På dansk ved Åge Haugland, Kbh.: NNF, Arnold Busck.
Lyons, J. 1981. Language and Linguistics. Cambridge University Press.
Macmillan, N.A., Kaplan, H.L. & Creelman, C.D. 1977. The psychophysics of categorical perception. PR 84: 452-471.
MacNeilage, P.F. 1970. Motor control of serial ordering of speech. PR 77: 182-196.
MacNeilage, P.F. 1972. Speech physiology. I Gilbert, J.H. (ed.): Speech and cortical functioning. 1-72.
MacNeilage, P.F. 1991. Comment: The gesture as a unit in speech perception theories. I Mattingly, I.G. & Studdert-Kennedy, M. (ed.): Modularity and the theory of speech perception. 61-68.
Madsen, T.K. 1996. Metaforer i den politiske diskurs. I Folia Scandinavica Posnaniensia 3: 273-284.
Massaro, D.W. 1970. Preperceptual Auditory Images. JEP 85(3): 411-417.
Massaro, D.W. 1971. Effect of Masking Tone Duration on Preperceptual Auditory Images. JEP 87(1): 146-148.
Massaro, D.W. 1972. Preperceptual Images, Processing Time, and Perceptual Units in Auditory Perception. PR 79(2): 124-145.
Massaro, D.W. 1975. Preperceptual Images, Processing Time, and Perceptual Units in Speech Perception. In Massaro, D.W. (ed.): Understanding Language. An Information-Processing Analysis of Speech Perception, Reading, and Psycholinguistics. 125-150. Academic Press.
Massaro, D.W. & Oden, G. C. 1980. Speech perception: A framework for research and theory. In Lass N.J. (ed.): Speech and Language: Advances in basic research and practice 3: 129-165. NY: Academic.
Mattingly, I.G., Liberman, A.M., Syrdal, A.M. & Halwes, T. 1971. Discrimination in speech and nonspeech modes. Cognitive Psychology 2: 131-157.
Mattingly, I.G. & Studdert-Kennedy, M. (ed.) 1991. Modularity and The Motor Theory of Speech Perception: Proceedings of a Conference to Honor Alvin M. Liberman Hillsdale, NJ: Lawrence Erlbaum Associates.
McGovern, K. & Strange, W. 1977. The perception of /r/ and /l/ in syllable-initial and syllable-final position. P&P 21: 162-170.
Mey, J.L. 1993. Pragmatics. An Introduction. Oxford: Blackwell Publishers.
Millar, J.M. & Whitaker, H.A. 1983. The right hemisphere's contribution to language. I Segalowitz, S.J. (ed.): Language functions and brain organization. 87-113. N.Y.: Academic Press.
Miller, C.L. & Morse, P.A. 1976. The heart of categorical speech discrimination in young infants. Journal of Speech and Hearing Research 19: 578-589.
Miller, G.A. 1956. The magical number seven, plus or minus two: Some limits in our capacity for processing information. PR 63: 81-97.
Miller, J.D., Wier, C.C., Pastore, R.E., Kelly, W.M. & Dooling, R.M. 1976. Discrimination and labeling of noise-buzz sequences with varying noise-lead times: An example of categorical perception. JASA 60: 410-417.
Miller, J.L. & Eimas, P.D. 1977. Studies on the perception of place and manner of articulation: A comparison of the labial-alveolar and nasal-stop distinction. JASA 61: 835-845.
Miyawaki, K., Strange, W., Verbrugge, R., Liberman, A.M., Jenkins, J.J. & Fujimura, O. 1975. An effect of linguistic experience: The discrimination of [r] and [l] by native speakers of Japanese and English. P&P 18: 331-340.
Morgan, A., McDonald, P. & MacDonald, H. 1971. Differences in bilabial alpha activity as a function of experimental task, with a note on lateral eye movements and hyponotizability. Neuropsychologia 9: 459-469.

Morris, C.W. 1946. Signs, Language and Behavior. Englewood Cliffs NJ: Prentice-Hall.
Morse, P.A. 1974. Infant speech perception: A preliminary model and review of the litterature. I Schiefelbusch, R. & Lloyd, L. (ed.): Language perspectives - acquisition, retardation, and intervention. Baltimore: University Park Press.
Morse, P.A. 1978. Infant speech perception: Origins, processes, and alpha centauri. I Proceedings of the NICHD Conference on "Early behavioral assessment of the communicative and cognitive abilities of the developmental disabled.
Morse, P.A. & Snowdon, C. 1975. An investigation of categorical speech discrimination by rhesus monkeys. P&P 17: 9-16.
Nebes, R.D. 1977. Menneskets såkaldt underordnede hemisfære. I Wittrock M.C. (ed.): Menneskets dobbelthjerne. En artikelsamling. 101-111. NNF.
O'Connor, J.D., Gerstman, L.J., Liberman, A.M., Delattre, P.C. & Cooper, F.S. 1957. Acoustic cues for the perception of initial /w, j, r, l/ in English. Word 13: 25-43.
Oden, G.C. & Massaro, D.W. 1978. Integration on feature information in speech perception. PR 85: 172-191.
Orr, D.B., Friedman, H.L. & Williams, J.C.C. 1965. Trainability of listening comprehension of speech discourse. Journal of Educational Psychology 56: 148-156.
Paradis, M. 1985. On the representation of two languages in one brain. Language Sciences 61: 1-39.
Paradis, M. 1995a. Another sighting of differential language laterality in multilinguals, this time in Loch Tok Pisin: Comments on Wuillemin, Richardson & Lynch (1994). Brain and Language 49: 173-186.
Paradis, M. 1995b. When the interpretation does not fit the facts, alter the facts, not the interpretation! A comment on Pulvermuller and Schumann (1994). Language Learning 45: 725-727.
Paradis, M. 1996. Selective deficit in one language is not a demonstration of different anatomical representation. Comments on Gomez-Tortosa et al. (1995). Brain and Language 54: 170-173.
Paradis, M. & Gopnik, M. 1994. Compensatory strategies in familial language impairment. McGill Working Papers in Linguistics 10: 142-149.
Pastore, R.E. 1981. Possible psychoacoustic factors in speech perception. In P.D. Eimas & J.L. Miller (eds.): Perspectives on the study of speech. Hillsdale, N.J.: Erlbaum.
Perey, A.J. & Pisoni, D.B. 1978. A reply to MacMillan, Kapland, and Creelman. Research on speech perception 4: 197-216.
Petersen, N.R. 1976. Identification and discrimination of vowel duration. ARIPUC 10: 57-84.
Petersen, N.R. 1995. Elementær statistik Institut for almen og anvendt sprogvidenskab, Kbh.
Peterson, G.E., Wang, W.S-Y. & Sivertson, E. 1958. Segmentation techniques in speech synthesis. JASA 30: 739-42.
Pisoni, D.B. 1971. On the nature of categorical perception of speech sounds. University of Michigan.
Pisoni, D.B. 1972. Perceptual Processing Time for Consonants and Vowels. Status Report on Speech Research SR-31/32: 83-92. Haskins Laboratories, New Haven.
Pisoni, D.B. 1973. Auditory and phonetic memory codes in the discrimination of consonants and vowels. P&P 13: 253-260.
Pisoni, D.B. 1975. The role of auditory short-term memory in vowel perception. Memory and Cognition 3: 7-18.
Pisoni, D.B. 1977. Identification and discrimination of the relative onset of two component tones: Implications for the perception of voicing in stops. JASA 61: 1352-1361.
Pisoni, D.B. & Lazarus, J.H. 1974. Categorical and non-categorical modes of speech perception along the voicing continuum. JASA 55: 328-333.
Pisoni, D.B. & Tash, J. 1974. Reaction times to comparisons within and across phonetic categories. P&P 15: 285-290.
Repp, B.H. 1975. Categorical perception, dichotic interference, and auditory memory: A "same-different" reaction time study. JASA 60: 456-469.
Repp, B.H. 1982. Phonetic Trading Relations and Context Effects: New Experimental Evidence for a Speech Mode of Perception. Psychological Bulletin 92(1): 81-110.
Repp, B.H. 1984. Categorical perception. Issues, methods, findings In N.J. Lass (ed.) Speech and Language: Advances in basic research and practice 10: 243-335. NY: Academic.
Repp, B.H., Healy, A.F. & Crowder, R.G. 1979. Categories and context in the perception of isolated steady-state vowels. JEP: Human Perception and Performance 5: 129-145.
Robbins, K. & McAdam, D. 1974. Interhemispheric alpha asymmetry and imagery mode. Brain and Language 2: 189-193.
Roberts, M. & Summerfield, Q. 1981. Audiovisual presentation demonstrates that selective adaptation in speech perception is purely auditory. P&P 30: 309-314.
Rosch, E. 1981. Prototype classification and logical classification: The two systems. I Scholnick, E. (ed.): New trends in cognitive representation: Challenges to Piaget's Theory, 73-86. Hillsdale, N.J.: Lawrence Erlbaum Associates.
Sawusch, J.R. & Jusczyk, P. 1981. Adaptation and contrast in the perception of voicing. JEP: Human Perception and Performance 7: 408-421.

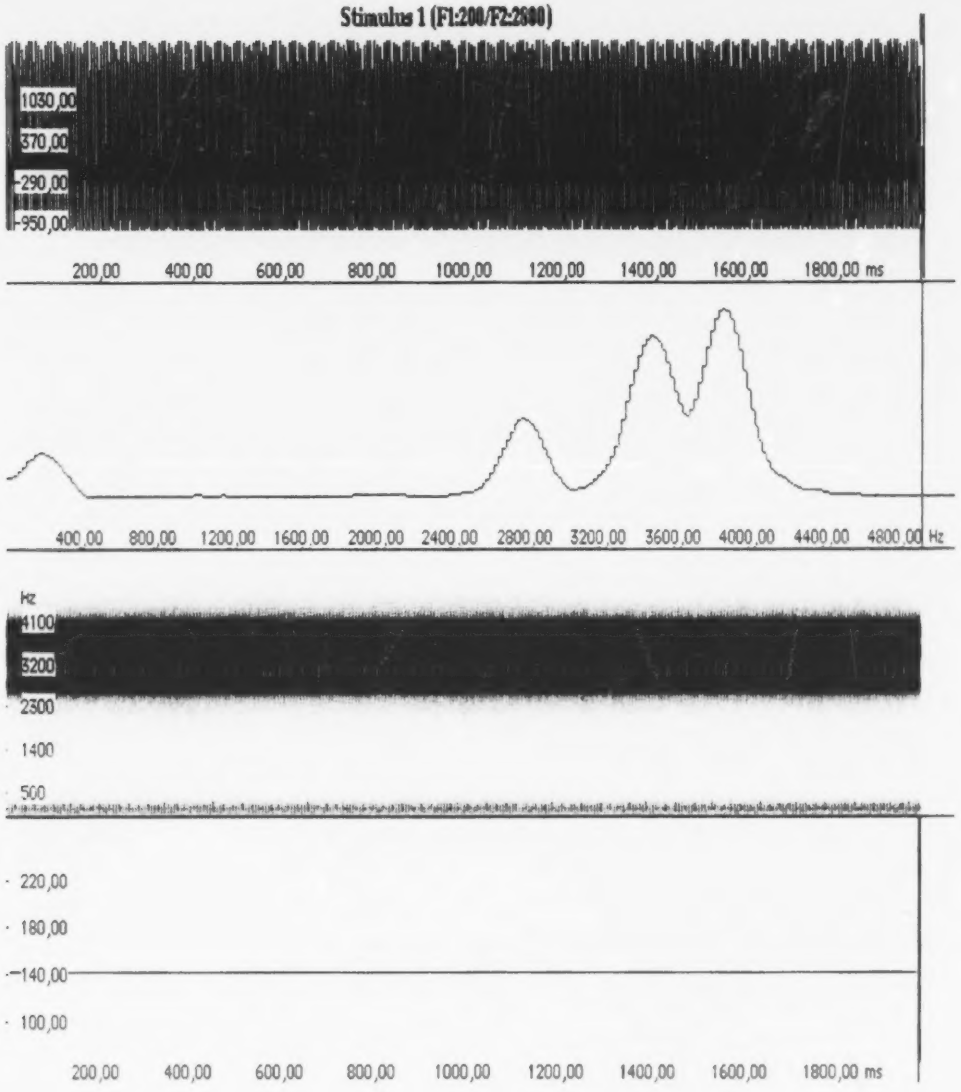
Schatz, C.D. 1954. The role of context in the perception of stops. Language 30: 47-56.
Schiffrin, D. 1994. Approaches to Discourse. Cambridge MA: Blackwell Publishers.
Schouten, M.E.H. 1980. The case against a speech mode of perception. Acta Psychologica 44: 71-98.
Schouten, M.E.H. (ed.). 1987. The Psychophysics of Speech Perception. Dordrecht: Martinus Nijhoff Publishers.
Segalowitz, S.J. & Gruber, F.A. (ed.). 1977. Language Development and Neurological Theory. New York: Academic Press.
Shankweiler, D.P. & Studdert-Kennedy, M. 1967. Identification of consonants and vowels presented to left and right ears. JEP 19: 59-63.
Sinnot, J.M., Beecher, M.D., Moody. D.B. & Stebbins, W.C. 1976. Speech sound discrimination by monkeys and humans. JASA 60: 687-695.
Spellacy, F. 1970. Lateral preferences in the identification of patterned stimuli. JASA 47: 574-578.
Sperber, D. & Wilson, D. 1986. Relevance. Communication and Cognition. Oxford: Blackwell Publishers.
Springer, S. 1971. Ear asymmetry in a dichotic detection task. P&P 10: 239-241.
Springer, S. 1979. Speech perception and the biology of language. I Gazzaniga M.S. (ed.): Handbook of Behavioral Neurobiology, vol. 2: 153-176, Neuropsychology.
Steinberg, D.D. 1993. An Introduction to Psycholinguistics. London: Longman Group.
Stevens, K.N. 1960. Toward a model of speech recognition. JASA 32: 47-55.
Stevens, K.N. & Halle, M. 1967. Remarks on analysis by synthesis and distinctive features. I Wathen-Dunn, W. (ed.): Models for the perception of speech and visual form. 88-102.
Stevens, K.N. & House, A.S. 1956. Studies of formant transitions using a vocal tract analog. JASA 28: 578-585.
Stevens, K.N., Liberman, A.M., Öhman, S.E.G. & Studdert-Kennedy, M. 1969. Crosslanguage study of vowel perception. Language and Speech 12: 1-23.
Strange, W. & Jenkins, J.J. 1978. Role of linguistic experience in the perception of speech. I Walk, R.D. & Pick, Jr. (ed.): Perception and experience. N.Y.: Plenum. 125-169.
Streeter, L. 1976. Language perception of two month old infants shows effects of both innate mechanisms and experience. Nature 259: 39-41.
Studdert-Kennedy, M. 1974. The perception of speech. I Seabeok, T.A. (ed.): Current trends in Linguistics, vol. 12. The Hague: Mouton.
Studdert-Kennedy, M. 1982. A note on the biology of speech perception. I Mehler, J. & Walker, E. (ed.): Perspectives in mental representation. Hillsdale, N.J.: Erlbaum.
Studdert-Kennedy, M., Liberman, A.M., Harris, K.S. & Cooper, F.S. 1970. Motor theory of speech perception: A reply to Lane's critical review. PR 1970: 234-249.
Studdert-Kennedy, M. & Shankweiler, D.P. 1970. Hemispheric specialization for speech perception. JASA 48: 579-594.
Teyler, T.J. 1977. Introduktion til neurovidenskaberne. I Wittrock, M.C. (ed.): Menneskets dobbelthjerne. En artikelsamling. 13-46. NNF.
Thorsen, N. & Thorsen, O. 1976. Lærebog i fonetik. Kbh.: Institut for fonetik.
Wada, J. & Rasmussen, T.R. 1960. Intracarotid injection of sodium amytal for the lateralization of cerebral dominance: Experimental and clinical observations. Journal of Neurosurgery 17: 266-282.
Warren, R.M. 1983. Multiple Meanings of "Phoneme" (Articulatory, Acoustic, Perceptual, Graphemic) and Their Confusions. I Lass N.J. (ed.): Speech and Language: Advances in basic research and practice 9: 285-312. NY: Academic.
Waters, R. & Wilson, W. 1976. Speech perception by rhesus monkeys: The voicing distinction in synthesized labial and velar stop consonants. P&P 19: 285-289.
Whitfield, H.A. & Evans, E.F. 1965. Responses of auditory cortical neurons to stimuli of changing frequency. Journal of Neurophysiology 28: 655-672.
Wittrock, M.C. (ed.). 1977. Menneskets dobbelthjerne. En artikelsamling. NNF, Prentice Hall Inc.
Wood, C.C. 1975. Auditory and Phonetic Levels of Processing in Speech Perception: Neurophysiological and Information-Processing Analyses. JEP. Human Perception and Performance 104(1): 3-20.
Wood, C.C., Goff, W. & Day, R. 1971. Auditory evoked potentials during speech perception. Science 173: 1248-1251.
Zurif, E. & Sait, P. 1969. The role of syntax in dichotic listening. Neuropsychologia 8: 239-244.

Forkortelser:

ARIPUC	Annual Report of the Institute of Phonetics, University of Copenhagen
JASA	Journal of the Acoustical Society of America
JEP	Journal of Experimental Psychology
P&P	Perception & Psychophysics
PR	Psychological Review

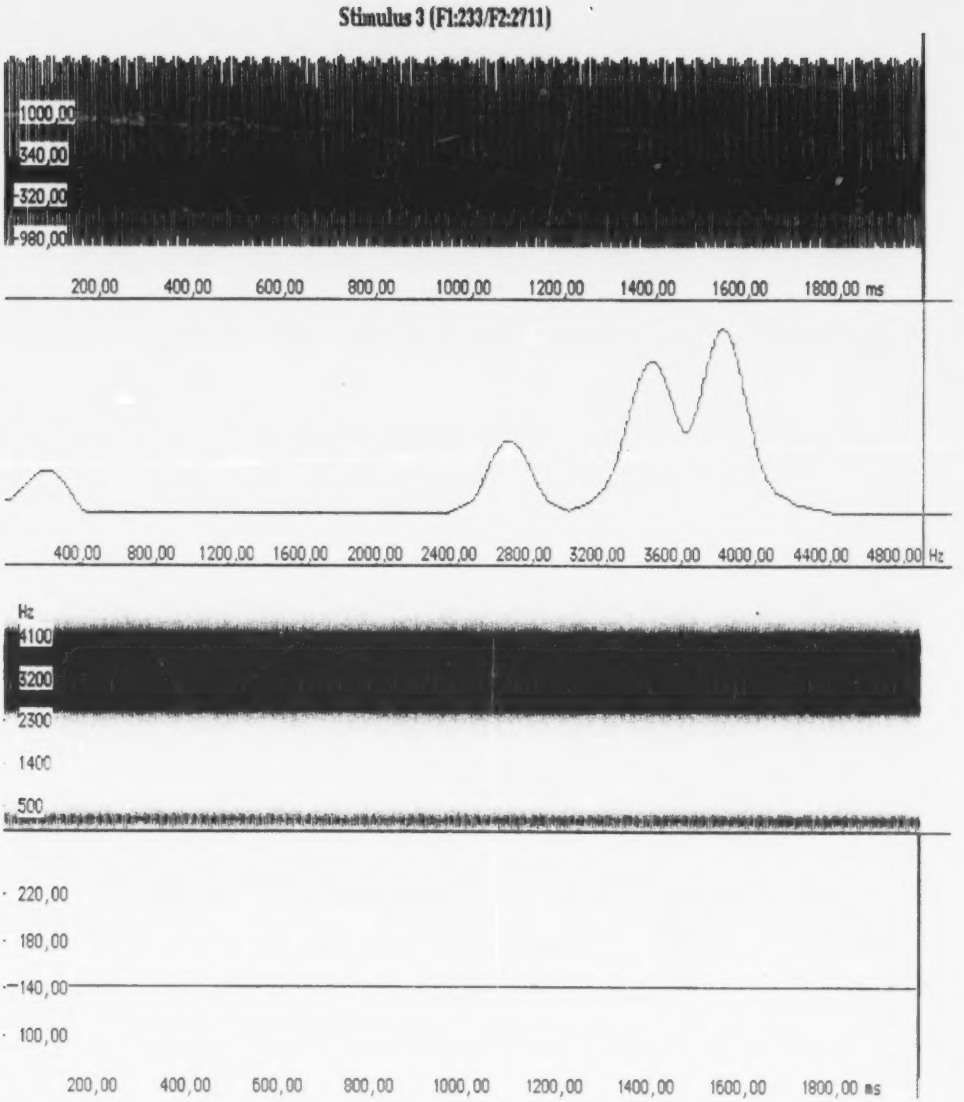
163

Bilag A: Akustiske analyser af syntetisk vokalkontinuum.



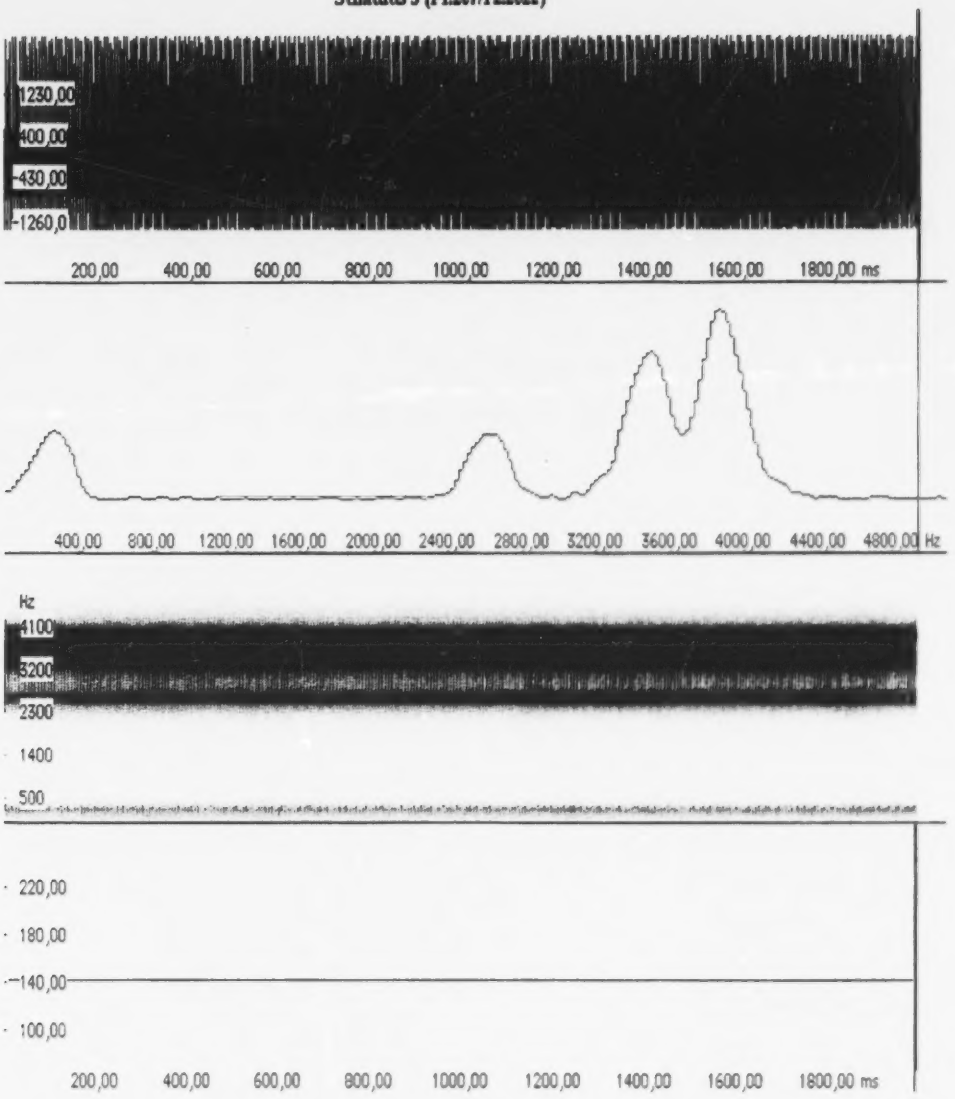
Vindue 1: Oscillogram
Vindue 2: FFT-Spektrum (0 - 5512 Hz)
Vindue 3: FFT-spektrogram (wide band, 8 ms/125 Hz)
Vindue 4: F0/grundtonen

BEST COPY AVAILABLE

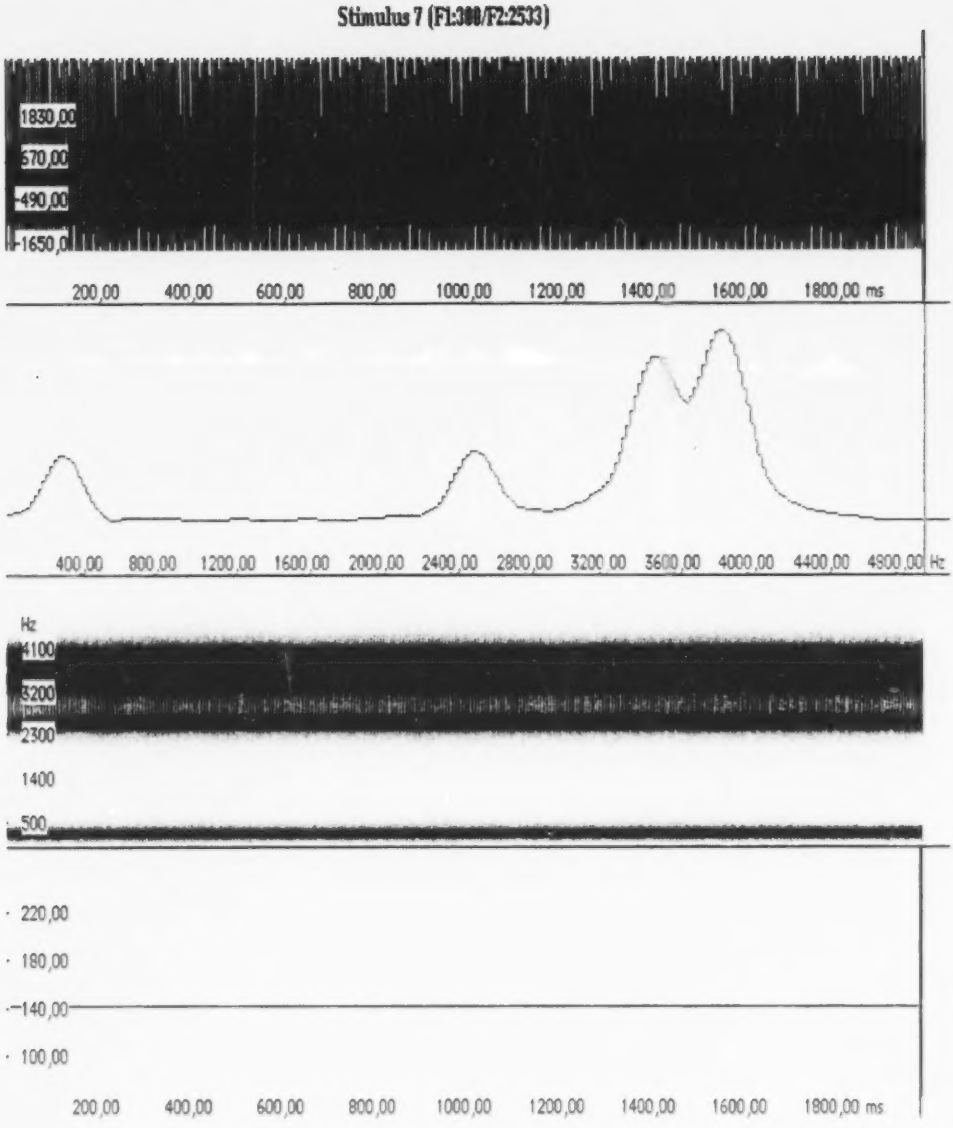


Vindue 1: Oscillogram
Vindue 2: FFT-Spektrum (0 - 5512 Hz)
Vindue 3: FFT-spektrogram (wide band, 8 ms/125 Hz)
Vindue 4: F0/grundtonen

170

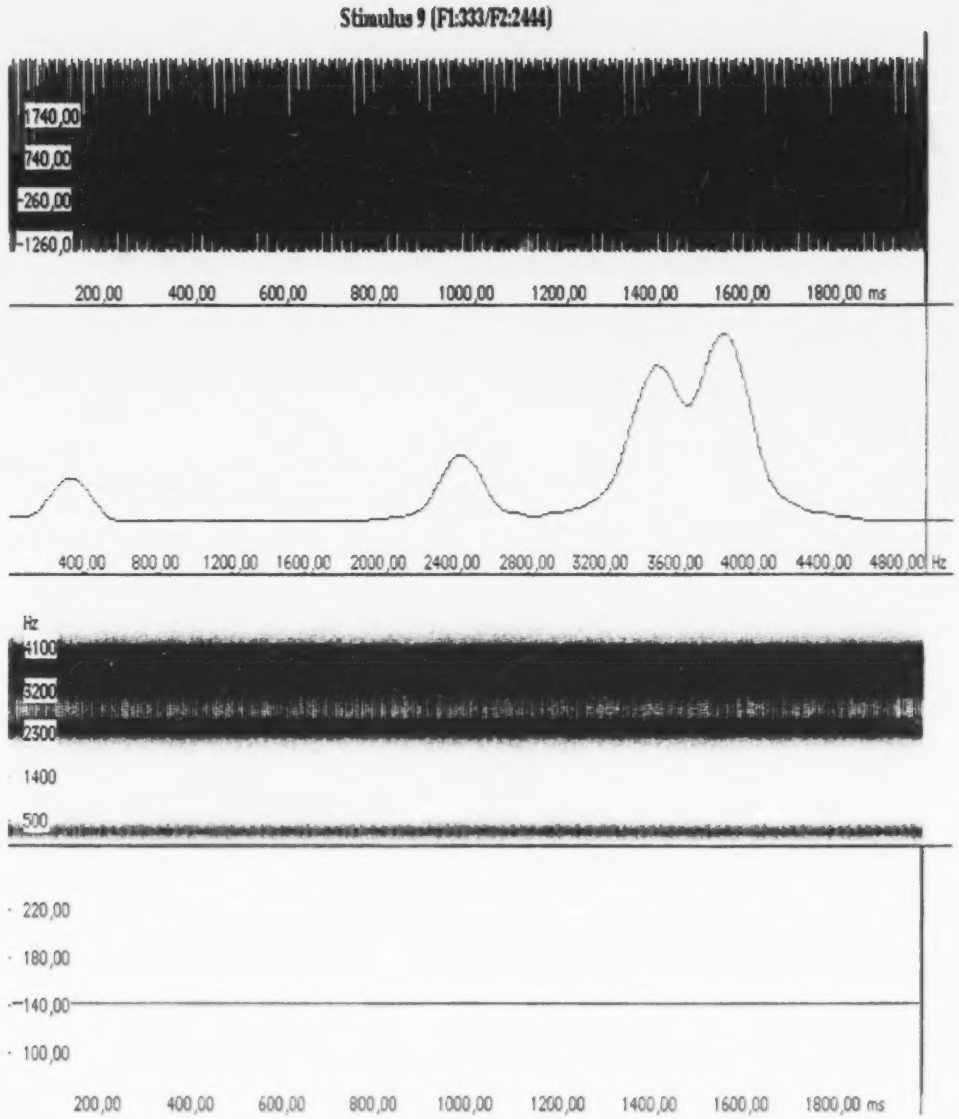


Vindue 1: Oscillogram
Vindue 2: FFT-Spektrum (0 - 5512 Hz)
Vindue 3: FFT-spektrogram (wide band, 8 ms/125 Hz)
Vindue 4: F0/grundtonen



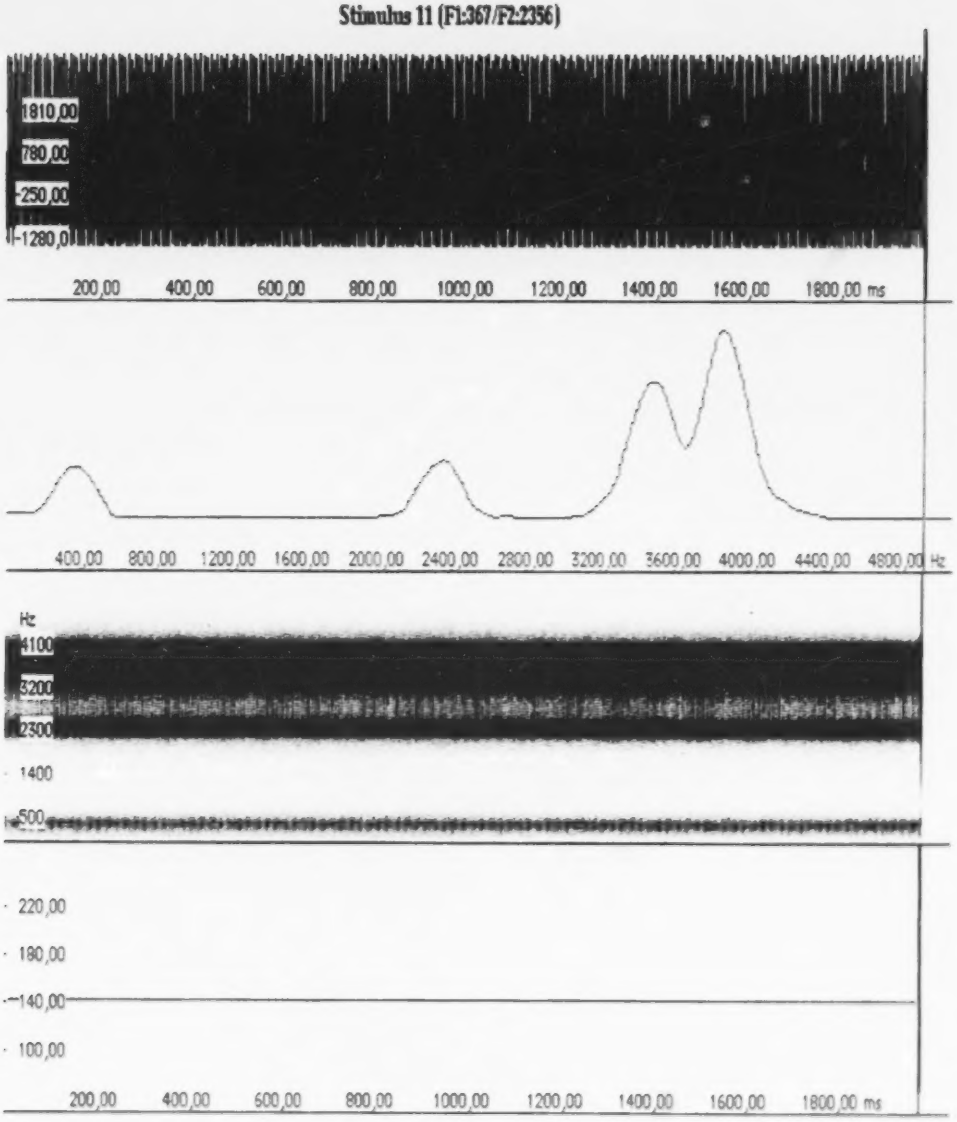
Vindue 1: Oscillogram
Vindue 2: FFT-Spektrum (0 - 5512 Hz)
Vindue 3: FFT-spektrogram (wide band, 8 ms/125 Hz)
Vindue 4: F0/grundtonen

172

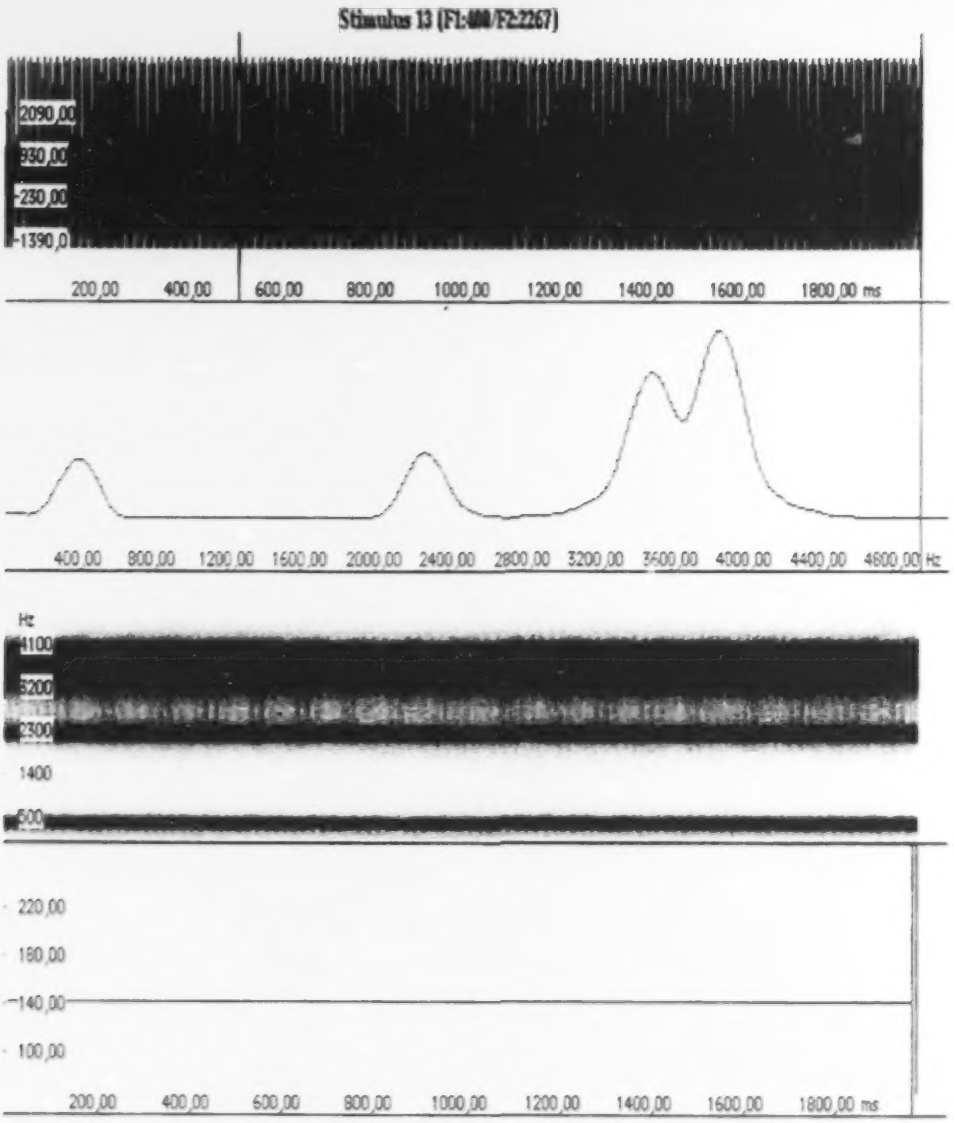


Vindue 1: Oscillogram
Vindue 2: FFT-Spektrum (0 - 5512 Hz)
Vindue 3: FFT-spektrogram (wide band, 8 ms/125 Hz)
Vindue 4: F0/grundtonen

BEST COPY AVAILABLE

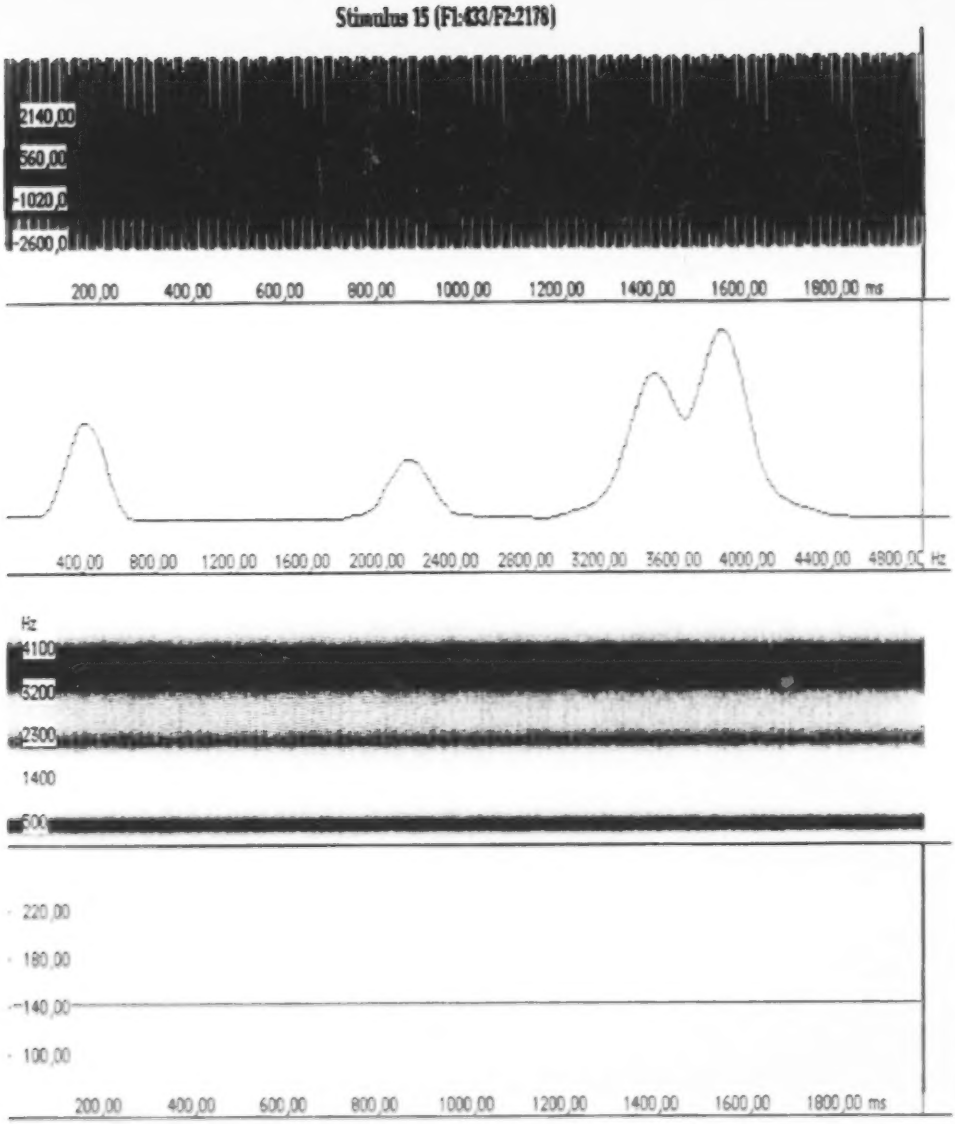


Vindue 1: Oscillogram
Vindue 2: FFT-Spektrum (0 - 5512 Hz)
Vindue 3: FFT-spektrogram (wide band, 8 ms/125 Hz)
Vindue 4: F0/grundtonen



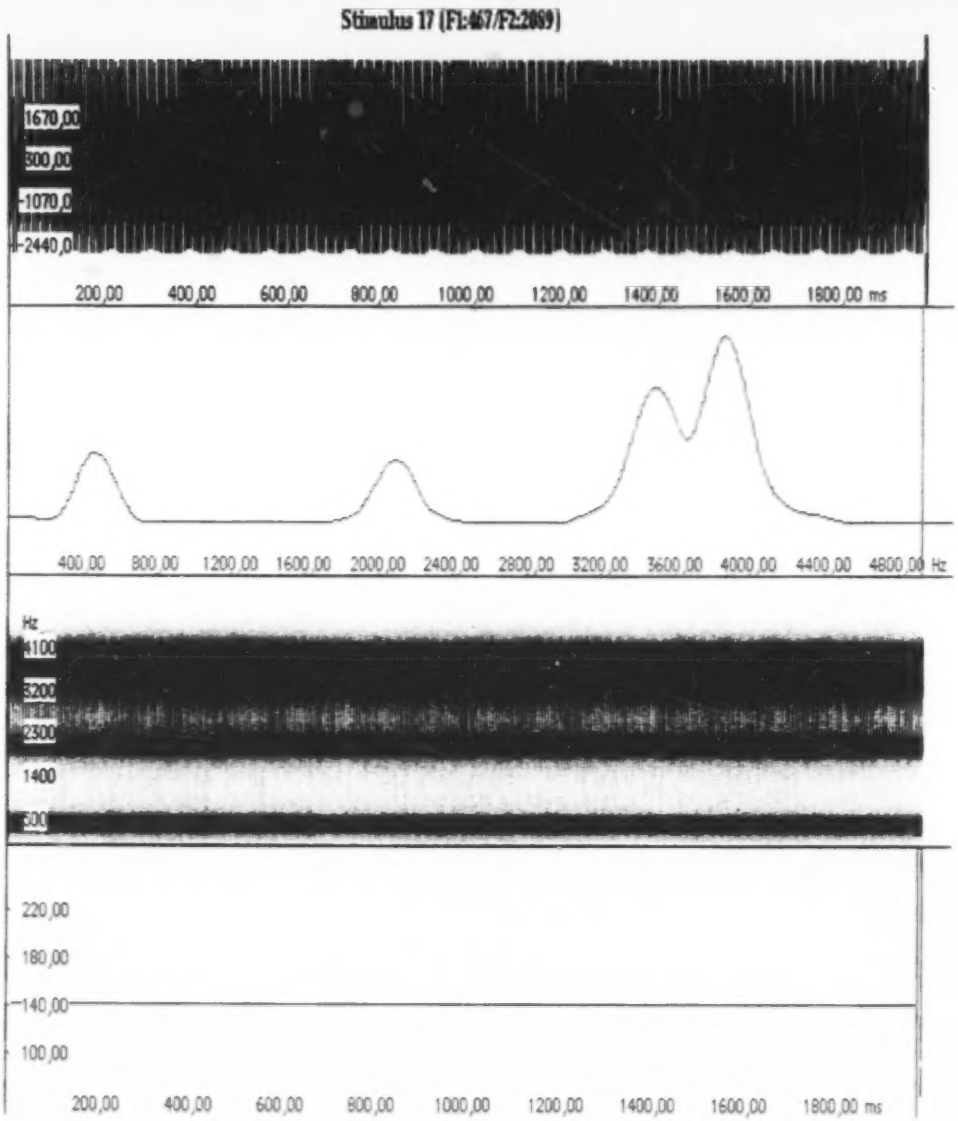
Vindue 1: Oscillogram
Vindue 2: FFT-Spektrum (0 - 5512 Hz)
Vindue 3: FFT-spektrogram (wide band, 8 ms/125 Hz)
Vindue 4: F0/grundtonen

BEST COPY AVAILABLE

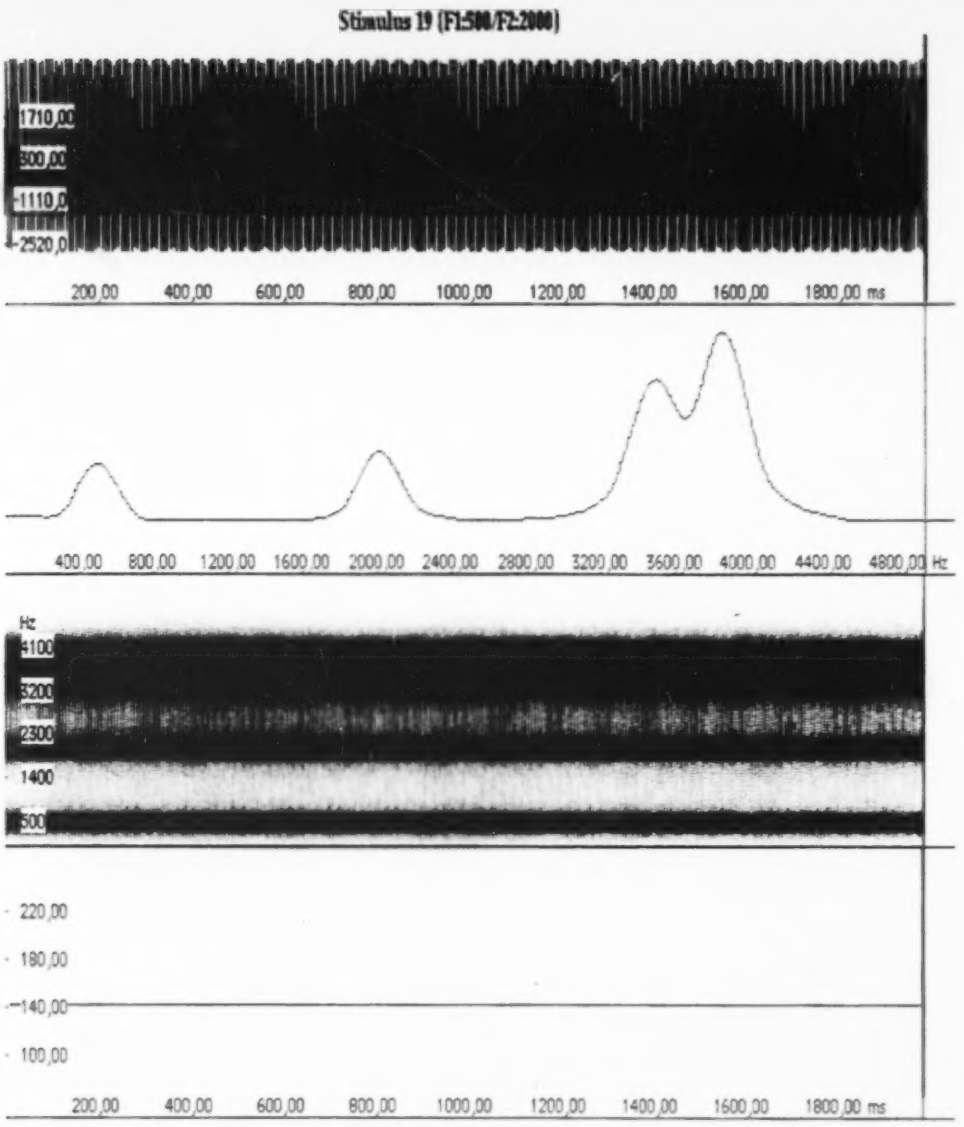


Vindue 1: Oscillogram
Vindue 2: FFT-Spektrum (0 - 5512 Hz)
Vindue 3: FFT-spektrogram (wide band, 8 ms/125 Hz)
Vindue 4: F0/grundtonen

17 0



Vindue 1: Oscillogram
Vindue 2: FFT-Spektrum (0 - 5512 Hz)
Vindue 3: FFT-spektrogram (wide band, 8 ms/125 Hz)
Vindue 4: F0/grundtonen



Vindue 1: Oscillogram
Vindue 2: FFT-Spektrum (0 - 5512 Hz)
Vindue 3: FFT-spektrogram (wide band, 8 ms/125 Hz)
Vindue 4: F0/grundtonen

178

WORD OF MOUTH

Communication is power. An idea, passed from person to person, and village to village, can transform the world.

Start with the right idea.

Linguistics and Language Behavior Abstracts
offering

- abstracts of scholarly articles and books
- bibliographic entries for subject specific dissertations and book and other media reviews.

LLBA's timely and comprehensive coverage speaks volumes on current ideas in linguistics and language research.

Available in print or electronically through the Internet Database Service (*www.csa.com*). Contact *sales@csa.com* for a trial Internet access or a sample issue.

LLBa

Linguistics and Language Behavior Abstracts
Published by Cambridge Scientific Abstracts
7200 Wisconsin Avenue • Bethesda, Maryland 20814 • USA
+1 301-961-6700 • Fax: +1 301-961-6720 • E-mail: *sales@csa.com*

BEST COPY AVAILABLE

Odense Working Papers in Language and Communication

OWPLC publishes scholarly articles written by members of the Institute of Language and Communication at the University of Southern Denmark, Main Campus: Odense University and their Danish or foreign colleagues.

Nos 14-20 include:

No. 14 Karl-Heinz Pogner (ed.):
Writing: Text and interaction

No. 15 Karin Van Durme (ed.):
The valency of nouns

No. 16 Dorthe Bleses & Johannes Wagner (eds.):
Papers in First Language Acquisition

No. 17 Dorthe Bleses:
The role of input, productivity and transparency in Danish children's acquisition of past tense morphology

No. 18 Teresa Cadierno (ed.):
L2 Listening comprehension

No. 19 Carl-Erik Lindberg & Steffen Nordahl Lund (eds.):
17th Scandinavian Conference of Linguistics, vol I-II 20-22 August 1998

No. 20 Thomas O. Madsen:
Sproglyde: fysiske kontinua og mentale kategorier

Copies can be obtained from

Institute of Language and Communication
University of Southern Denmark,
Main Campus: Odense University
Campusvej 55
DK-5230 Odense M.
Denmark

General Editor: Steffen Nordahl Lund
(E-mail: snl@language.sdu.dk.)

130

Odense Working Papers in Language and Communication publishes scholarly articles written by members of the Institute of Language and Communication at SDU - Odense University as well as papers by their Danish or foreign colleagues.

This study presents an empirical investigation of basic processes in the perception of speech sounds. The experimental methods applied have their roots in two different psycholinguistic research paradigms, i.e. 'categorical perception' and 'dichotic listening'.

In the categorical perception paradigm, listeners' categorization of speech sounds is compared with their ability to discriminate speech-like stimuli forming a physical continuum (such as the range between a t-sound and a d-sound).

In the dichotic listening paradigm, differences in listeners' identification of speech stimuli presented to either left or right ear are interpreted as the result of hemispheric lateralization.

BEST COPY AVAILABLE